# DEVELOPMENT AND APPLICATION OF A MICROBIAL RELIABILITY MODEL TO ANALYZE DENITRIFYING BIOFILTER STABILITY

BY

NICHOLAS A. BARTOLERIO

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Environmental Engineering in Civil Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Advisers:

Research Assistant Professor Julie L. Zilles
Assistant Professor Luis Rodríguez

# Abstract

The use of nitrogenous fertilizers to amend soil fertility has had a drastic effect on the global nitrogen cycle as excess nitrogen has contaminated surface waters, leading to marine hypoxic zones. One of the largest contributors to nitrogen loads in surface waterways is subsurface agricultural drainage. For this reason, efforts are underway to reduce drainage nitrogen losses to waterways. One proven method is the use of edge-of-field denitrifying biofilters. These biofilters redirect tile drainage flow through a woodchip bed, where microbial activity converts influent nitrate to nitrogen gases.

As with many engineered ecosystems, performance stability of the biofilters is of concern. While much research has focused on stability concepts in ecology, there is no consensus on the nature of the relationships between microbial diversity, functional redundancy, and ecosystem stability. A better understanding of these relationships and other factors that affect performance stability of the denitrifying biofilters is necessary to further enhance their effectiveness. With the uncertainties that currently exist in the area of ecosystem stability in mind, it was my goal to develop methods using reliability theory as a new approach to analyze ecosystem stability, linking the stability of microbial populations to the stability of the system functional performance.

To apply reliability theory to engineered ecosystems, microbial populations were represented as components in a system. We have developed a method of utilizing microbial fingerprinting data to quantify presence and longevity of microbial populations from which a reliability function for each population can be determined. We were able to target functional genes and identify populations directly responsible for the system functional performance using microbial fingerprinting techniques such as terminal restriction

fragment length polymorphism (T-RFLP). This allowed us to get a better understanding of how to model the microbial populations as functional components in the system.

In order to quickly and easily apply these methods to microbial fingerprinting data, we have developed the Ecosystem Reliability Analysis Tool (EcoReliAnT) in MATLAB®. This tool provides the functionality necessary for ecosystem reliability analysis through a user-friendly interface.

The newly developed reliability analysis methods and EcoReliAnT software was first tested on a dataset acquired from the literature. This external dataset consisted of phenol degradation rate performance data from a sequencing batch reactor along with corresponding microbial functional gene information from restriction fragment length polymorphism analysis. This dataset provided a good trial dataset for the methods and software, as it showed clear changes in microbial community structure that corresponded to changes in system performance.

Following the trial run on the external dataset, data from a field denitrifying biofilter in Decatur, Illinois was analyzed. This data consisted of *nosZ* -T-RFLP microbial community information from a 135 day period in which there was continuous flow along with percent nitrate removal as a performance metric.

Results from the reliability analyses of both the external dataset and the data from the field denitrifying biofilter demonstrated the capabilities of the reliability methods and EcoReliAnT software in analyzing engineered ecosystems. It was demonstrated that reliability functions could be determined for microbial populations based on their population dynamics and that the reliability of the system performance could be accurately modeled as a configuration of functional microbial components, with sum of squared error

values between reliability functions of the system and model as low as 0.07. In both cases, incorporating microbial populations that were determined to be negatively correlated with system performance along with those that were positively correlated with system performance resulted in the best-fitting model. This suggests that 'nuisance' populations play an important role in the stability of engineered ecosystems.

In addition to the development of reliability analysis methods, laboratory-scale denitrifying biofilters were designed and built to allow for a more controlled study of factors that affect system performance. These laboratory biofilters were fed synthetic tile drainage and allowed examination of the effects that changes in environmental and operational conditions have on the microbial community and system performance.

Results from the startup of the laboratory-scale biofilters suggest that under constant environmental and operational conditions, denitrifying biofilters can exhibit very stable performance. This highlights the importance of understanding how changes in these conditions affect both the system performance and the microbial community structure in order to better understand the performance stability of the system. The successful startup of the laboratory-scale biofilters provides a platform for future experiments to enhance the understanding of how microbial population dynamics, environmental parameters, and operational conditions relate to biofilter performance and stability.

The application of reliability theory methods to engineered ecosystems is a unique approach in the consideration of microbial diversity, functional redundancy, and engineered ecosystem stability. Developing a better understanding of the relationships between these concepts and other factors that affect the functional stability of denitrifying

biofilters will allow for improved system design and operation, ultimately enhancing their efficacy as a treatment technology.

# Acknowledgements

I would like to express my sincere gratitude to my two advisors, Dr. Julie Zilles and Dr. Luis Rodríguez, for their support, knowledge, motivation, and enthusiasm, without which this thesis would not be possible. Dr. Zilles allowed me to pursue my interest in laboratory reactor studies on this project and provided outstanding guidance throughout my graduate work. Dr. Rodríguez exposed me to reliability engineering and computer programming, two areas in which I had little previous training but have become very well acquainted with over the past two years.

I would also like to express my gratitude to Dr. Angela Kent for allowing me to use her lab and for helping an engineer try to think like a microbial ecologist. I would like to thank Dr. J. Malia Andrus and Matt Porter for providing the microbial data for my reliability analysis, along with providing field and lab instruction. I'd also like to thank Matt Porter further for his help with field sampling and assistance in crafting the laboratory biofilters. Paul Choi for his help operating and maintaining the laboratory reactors. Jason Koval for training and help running GC samples. I'd also like to thank the members of the Zilles and Rodríguez groups for their thoughtful insight. And finally, Dr. Richard Cooke and his laboratory for allowing access to the field sites and providing system performance data.

Most importantly, I would like to thank my friends and family for their support, in particular my girlfriend Katie for her encouragement, understanding, and patience during the past two years. I couldn't have done it without you.

This work was funded by the National Science Foundation Environmental Sustainability program.

# Table of Contents

# Chapter 1: Introduction

Human activity has had a major impact on the global nitrogen cycle. Along with the combustion of fossil fuels that releases nitrogen oxides into the atmosphere, human activity has led to increased nitrate concentrations in surface waters. This is especially evident in the agriculturally dominated Midwestern United States.

Increasing global population and the desire for high crop yields has led to widespread use of nitrogenous fertilizers to amend soil fertility. Since the nitrogen in these fertilizers is not completely utilized for crop growth, excess exists that migrates into groundwater and streams via surface and subsurface runoff. Nitrate is especially mobile in the soil, as its negative charge does not facilitate strong adsorption to soil particles. The high mobility of nitrate coupled with the agricultural water drainage systems used throughout much of the Midwestern United States allows for rapid transport of fertilizer nitrogen to local waterways.

Although local wastewater treatment facilities discharge nitrate-containing effluent into the same waterways, inputs from wastewater are estimated to make up only 9% of the nitrate load (Howarth et al. 1996). By far the largest contributor of riverine nitrate in the Mississippi River Basin is subsurface runoff from fertilized cropland with tile drainage systems (Goolsby et al. 1999; David et al. 2010). In Illinois, drainage water nitrate losses have been found to range from 23-33 kg $NO_3$-N/ha/yr and had average concentrations of 15-20 mg/L $NO_3$-N, which is above the EPA drinking water limit (Kalita et al. 2006).

Increased nitrate loads in waterways can have drastic ecological effects. One major effect is the increase in primary production in previously nitrate-limited systems, leading

to enhanced algal growth in receiving waters. Algal growth and subsequent decomposition deplete dissolved oxygen, creating hypoxic zones. These areas of less than 2 parts per million of dissolved oxygen, including a large zone in the Northern Gulf of Mexico, are no longer able to sustain marine life. The number of hypoxic zones in coastal ocean areas has increased exponentially since the 1960s, and there are currently over 400 globally (Diaz and Rosenberg 2008).

To curb nitrate losses to waterways, biological denitrification systems have been designed to remove nitrate from runoff. These systems include denitrification walls (Starr and Cherry 1994; Schipper and Vojvodic-Vukovic 1998; Robertson et al. 2000; Schipper and Vojvodic-Vukovic 2000; Schipper and Vojvodic-Vukovic 2001; Schipper et al. 2004; Schipper et al. 2005), constructed wetlands (Kovacic et al. 2000), riparian buffer strips (Yamada et al. 2007; Blattel et al. 2009; Woodward et al. 2009;), and denitrifying biofilters (Blowes et al. 1994; Cooke et al. 2001; Wildman 2002; Doheny 2003; van Driel et al. 2006; Jaynes et al. 2008; Chun et al. 2010; Schipper et al. 2010; Woli et al. 2010). The University of Illinois has installed several edge-of-field biofilters, consisting of trenches filled with woodchips through which the tile drainage flows (Cooke et al. 2001; Wildman 2002; Doheny 2003; Chun et al. 2010; Woli et al. 2010).

As with other engineered ecosystems, stable performance of the denitrifying biofilters is of interest for future design and operation. Although traditional engineering studies have been undertaken (Cooke et al. 2001; Wildman 2002; Doheny 2003; Chun et al. 2010), questions regarding inconsistent nitrate removal performance in the biofilters remain. This has led to recent work analyzing the microbial communities responsible for nitrate removal in the biofilters in an attempt to use microbial ecology to improve biofilter

functionality (Andrus 2010; Porter 2011). While microbial ecology provides unique insights into ecosystem community composition and population dynamics, uncertainties regarding the application of these methods for the analysis and improvement of ecosystem stability persist. Along with varying definitions that have been used for ecosystem stability (Grimm et al. 1992), inconsistent results have come from studies comparing microbial community stability with functional performance stability (Martienssen and Schops 1997; Fernandez et al. 1999; von Canstein et al. 2001; Smith et al. 2003; Stamper et al. 2003; Cytryn et al. 2005; Gentile et al. 2006; Miura et al. 2007; Wang et al. 2011). Likewise, studies examining the relationship between microbial diversity and functional stability have failed to reach a consensus (Daims et al. 2001; Rowan et al. 2003; Griffiths et al. 2004; Girvan et al. 2005; Pholchan et al. 2010). Clearly, further study into the relationship between microbial community and performance stability of engineered ecosystems is required. In this thesis, we propose the use of reliability theory as a novel approach that links the stability of microbial communities and system functional performance to quantify engineered ecosystem stability. Methods to apply reliability techniques to engineered ecosystems were developed, along with a new software tool, the Ecosystem Reliability Analysis Tool (EcoReliAnT).

The research presented in this thesis focuses on the development and application of our ecosystem reliability analysis methods and is comprised of three separate studies. The first study applies our reliability analysis approach to a relatively simple engineered ecosystem from the literature (Ayala-del-Rio et al. 2004). Ayala-del-Rio and colleagues analyzed microbial community population dynamics in phenol and phenol-plus-TCE fed sequencing batch reactors (2004). This dataset gave us an opportunity to apply our

analysis approach to a less complex system, as microbes possessing the phenol hydroxylase gene are less diverse than those in the denitrification pathway and the system is only comprised of one functional step.

The second study then applies the same approach to the microbial population data from a field denitrifying biofilter from central Illinois (Porter 2011). For the field biofilters, we have microbial population information for total bacterial (ARISA), denitrifying bacterial (*nosZ* T-RFLP), and fungal communities (FARISA). This is a much more complex system than the phenol-fed sequencing batch reactor system, with hundreds of microbial populations present and varying environmental and operational conditions.

The third study consists of the design and operation of laboratory-scale denitrifying biofilters seeded with woodchips from a field biofilter. These laboratory-scale biofilters allow us to closely monitor microbial community structure and biofilter performance while controlling many environmental and operational parameters such as hydraulic retention time, nitrate concentration, and water depth. These laboratory-scale biofilters will supply system performance and microbial community data for future reliability analysis while providing a platform for further study of the relationships between environmental and operational parameters, microbial diversity, and ecosystem stability.

I believe that these studies on the application of reliability theory methods to engineered ecosystems will enhance the understanding of the effect of microbial diversity and stability on ecosystem functional performance. This new approach to modeling microbial ecosystems should also enable the identification of components that limit overall system performance and identify avenues for ecosystem performance improvement in the denitrifying biofilters as well as other engineered ecosystems. A more comprehensive

understanding of the stability of denitrifying biofilters is necessary for the optimization of design and operation and will be a key component in the advancement of this technology for reducing nutrient loads in agricultural runoff.

# Chapter 2: Literature Review

To aid in the understanding of this research, this literature review provides background information relevant to our study on denitrifying biofilters and introduces fundamental concepts that will be used in our reliability analysis of engineered ecosystems. In this section, I will discuss five topics; (1) various engineered denitrification systems; (2) previous studies on denitrifying biofilters both at the field and laboratory scale; (3) basic concepts used to measure microbial ecosystems; (4) stability concepts in ecology; and (5) the methods used in system reliability analysis.

## 2.1 Engineered Denitrification Systems

Denitrification is the microbially mediated process that reduces nitrate to dinitrogen gas. Although microbes capable of denitrification have been identified in a variety of bacterial genera and display significant metabolic and physiological diversity, most denitrifiers are facultative heterotrophs (Zumft 1997). These denitrifiers utilize nitrate as an electron acceptor under anoxic conditions, requiring organic matter as an electron donor. This cellular respiration results in successive reduction of the nitrogen species, each of which is mediated by a specific enzyme (Figure 1). Any given denitrifying microorganism may possess the genes required to produce either a partial or complete set of these enzymes, and thus are capable of mediating some or all of these reactions.

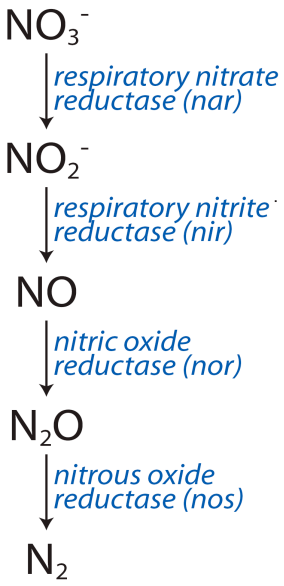



Figure 1. Denitrification pathway and the reductase enzymes (and genes) responsible for each step in the reduction of nitrate to dinitrogen gas.

Several different engineered denitrification systems have arisen for treating agricultural runoff in various hydrological conditions, including riparian buffer strips, denitrification walls, constructed wetlands, and denitrifying biofilters . One of the most important considerations in selecting a denitrification system is whether the agricultural land providing the runoff is artificially or naturally drained.

In agricultural land that is not tile drained, denitrification systems must intercept groundwater seepage before it enters surrounding waterways. One method used to remove nitrate from groundwater is the installation of riparian buffer strips consisting of a zone of dense hydrophilic plant growth between the agricultural fields and receiving waters (Yamada et al. 2007; Blattel et al. 2009; Woodward et al. 2009). These systems utilize microbial denitrification along with biological uptake to remove nitrate from groundwater. In a survey of 45 published studies including 88 riparian buffer strips, Mayer

et al. calculated a mean nitrate removal of 67.5+- 4.0% and determined that the largest factor affecting nitrate removal efficiency was the width of the buffer (2007).

Another method used to remove nitrate from laterally moving groundwater is the installation of a denitrification 'wall' consisting of a trench that has been excavated perpendicular to the direction of flow and replaced with a mixture of soil and sawdust (Starr and Cherry 1994; Schipper and Vojvodic-Vukovic 1998; Robertson et al. 2000; Schipper and Vojvodic-Vukovic 2000; Schipper and Vojvodic-Vukovic 2001; Schipper et al. 2004; Schipper et al. 2005). One study found nitrate removal rates of 1.4 g N $m^{-3}$ of wall $d^{-1}$ in a denitrification wall treating groundwater dosed with nitrate (Schipper et al. 2005). This is similar to values seen in other studies (Robertson et al. 2000), although short-circuiting underneath the denitrification wall has also been a concern (Schipper et al. 2004).

In artificially drained fields, including much of the Midwestern United States (Jaynes and James 2007), drainage systems move runoff from a large field area through a network of pipes that ultimately converge to a single outfall. This piping system essentially provides a point source, allowing for a treatment system at a single location.

Constructed wetlands installed at the end of tile drainage lines have demonstrated 31-42% nitrate removal (Kovacic et al. 2000; Kovacic et al. 2006). Although constructed wetlands have been shown to remove over 90% of nitrate in other applications, achieving this level of removal requires large areas, warm temperatures, and long retention times (>1 day)(Reviewed in O'Geen et al. 2010). The large size of wetlands required to treat runoff not only take away potential cropland, but also make constructed wetlands relatively

expensive to install (O'Geen et al. 2010). Because of this farmers may be reluctant to install constructed wetlands to treat their tile drainage..

Another technology utilized to remove nitrate from tile drainage is the edge-of-field denitrifying biofilter (Cooke et al. 2001; Wildman 2002; Doheny 2003; Chun et al. 2010; Woli et al. 2010). Denitrifying biofilters are an effective, low-maintenance, and low-cost denitrification system and are the focus of this thesis. These biofilters will be discussed in detail in the following section.

## 2.2 Denitrifying Biofilters

A cost-effective denitrification system that has successfully removed nitrate from drainage water is the edge-of-field denitrifying biofilter. The University of Illinois has installed several of these biofilters throughout central Illinois (Cooke et al. 2001; Wildman 2002; Doheny 2003; Chun et al. 2010; Woli et al. 2010). Denitrifying biofilters consist of trenches filled with woodchips through which the tile drainage flows. Denitrifying microorganisms from the soil inoculate the woodchip beds, using organic carbon from the microbial degradation of the woodchips as an electron donor and aqueous nitrate as an electron acceptor. Denitrifying biofilters are effective in removing nitrate from drainage water. In a two-year field study of a bioreactor in Illinois, nitrate reduction efficiencies ranged from 12% to 99.5% with an overall nitrate load removal of 33% (Woli et al. 2010). Similar denitrifying biofilters have also been installed in other locations (Blowes et al. 1994; van Driel et al. 2006; Jaynes et al. 2008; Schipper et al. 2010).

Nitrate removal rates of 0.9-2.5g $N/m^2/d$ shown in denitrifying biofilters are an order of magnitude larger than those typically seen in constructed wetlands (van Driel et

al. 2006). Therefore, denitrifying biofilters require a land investment one-tenth that of wetlands (Kovacic et al. 2000). Furthermore, recent biofilter designs have the woodchip bed covered with approximately two feet of topsoil, allowing crops to be grown on top of the biofilter. This minimizes the required land investment and alleviates the concern of losing valuable cropland.

### 2.2.1 Field Biofilter Studies

Several studies have focused on improving the design of denitrifying biofilters to optimize performance or minimize cost and maintenance (Wildman 2002; Chun et al. 2010; Moorman et al. 2010). In a study of field biofilters in central Illinois, 95-100% nitrate removal efficiencies occurred after a few month startup period of 3-11% removal, suggesting a fairly rapid enrichment of the microbial community with denitrifiers (Wildman 2002). A further study on one of these biofilters examined flow and transport parameters, estimating the longitudinal and transverse dispersivity (10.2 and 1.13 cm), effective porosity (0.79), and first order decay coefficient ($0.01 h^{-1}$) (Chun et al. 2010). Another study examined denitrification activity, wood loss, and nitrous oxide production in field biofilters over a 9-year span (Moorman et al. 2010). Denitrification potential rates ranged from 8.2 to 34.4 mg N $kg^{-1}$ wood $d^{-1}$, matching the estimated nitrate removal rate in the field of 23.6 mg N $kg^{-1}$ wood $d^{-1}$. Dissolved nitrous oxide exports in the treated tile drainage were not significantly higher than exports in non-treated drainage. Woodchips that were permanently submerged decomposed at a slower rate (20% loss versus 75% loss) than those that were intermittently submerged. These field studies have shown that denitrifying biofilters are a cost-effective and low maintenance method to remove nitrate from agricultural runoff. They have also suggested that complete denitrification occurs in

the biofilters and therefore nitrous oxide emissions should not be a concern for the adoption of this treatment technology.

### 2.2.2 Lab Biofilter Studies

Along with studies of denitrifying biofilters in the field, there have been many studies undertaken at the laboratory-scale. One focus has been on different potential carbon sources for denitrification (Volokita et al. 1996; Ines et al. 1998; Doheny 2003; Greenan et al. 2006; Cameron and Schipper 2010; Gibert et al. 2008). These studies demonstrated that while many different carbon sources can be utilized in denitrification systems including wheat straw (Ines et al. 1998; Cameron and Schipper 2010;, cotton (Volokita et al. 1996), and maize cobs (Cameron and Schipper 2010), woodchips are ideal for use in denitrification beds because of their low cost, high hydraulic conductivity, and sustained nitrate removal (Volokita et al. 1996; Doheny 2003; Greenan et al. 2006; Gibert et al. 2008; Cameron and Schipper 2010). A study using PVC pipe reactors filled with woodchips concluded that nitrate transport was best explained by a first-order reaction (Chun et al. 2009). Another study found that denitrification rates increased as flow rate increased in laboratory biofilters (Greenan et al. 2009). Denitrification was found to be the dominant nitrogen removal mechanism. Nitrous oxide production was 0.003 to 0.028% of the N denitrified, indicating that complete denitrification occurred. Denitrifying biofilters have also been studied at the pilot scale to examine the hydraulic properties of biofilter media (Christianson et al. 2010a) and the effects of reactor geometry on performance at different hydraulic retention times, finding a linear increase in nitrate removal from 30 to 70% with retention times of 4 to 8 hours (Christianson et al. 2010b; Christianson et al. 2011).

### 2.2.3 Microbial Community Studies on Denitrifying Biofilters

The previously discussed studies, both at field and laboratory scale, have examined many questions regarding denitrifying biofilters from a traditional engineering perspective. Although the performance of the denitrifying biofilters relies on microbial activity, only recently has the microbial ecology of the biofilters been examined (Andrus 2010; Porter 2011). Andrus analyzed both total and *nosZ*-possessing bacteria for spatial and temporal patterns in biofilters in central Illinois (Andrus 2010). The total bacterial community showed spatial structure over a single biofilter but the denitrifying community did not. Environmental parameters such as inlet nitrate, port temperature, moisture content, and pH were also associated with variation in community composition. Studying the same biofilter and two others over time, Porter found that total bacterial, fungal, and denitrifying communities were structured by season and depth in the biofilter (Porter 2011). This suggests that temperature or moisture content gradient may drive the community structure. These studies enhanced understanding of the microbial communities in the biofilters, but the relationship between microbial community structure and system performance stability is still uncertain.

## 2.3 Microbial Community Measurements

"Microbial fingerprinting" is a term used to describe molecular approaches to distinguish populations within a microbial community. Modern high throughput techniques utilizing polymerase chain reaction (PCR) have given researchers the means to

rapidly create these 'fingerprints' and examine microbial diversity and changes in microbial community structure. There are several methods available to create a 'fingerprint' of a microbial community in a sample, including denaturing gradient gel electrophoresis (DGGE) (Muyzer et al. 1993), Amplified rDNA Restriction Analysis (ARDRA)(Vaneechoutte et al. 1993), Automated Ribosomal Intergenic Spacer Analysis (ARISA)(Fisher and Triplett 1999), Restriction Fragment Length Polymorphism (RFLP)(Moyer et al. 1994), and Terminal Restriction Fragment Length Polymorphism (T-RFLP) (Liu et al. 1997). Many fingerprinting techniques involve the 16S ribosomal RNA (rRNA), as this component is highly conserved across bacteria (Woese 1987). The 16S component of rRNA is part of the 30S subunit of a ribosome and is involved in the production of proteins. The importance of this cell function has led to the high conservation of the 16S rRNA gene, making it useful for phylogenetic studies. In the ecosystem reliability analysis in this thesis, RFLP and T-RFLP are employed to examine microbial community structure and will be discussed in more detail in the following sections.

### 2.3.1 RFLP and T-RFLP

Restriction Fragment Length Polymorphism (RFLP)(Moyer et al. 1994) and Terminal Restriction Fragment Length Polymorphism (T-RFLP) (Liu et al. 1997) are molecular fingerprinting techniques that use gene-specific PCR primers to profile microbial populations based on restriction fragment lengths. Nucleic acid PCR products are digested with restriction enzymes that cut the nucleic acid into fragments of varied lengths. These restriction fragment lengths depend upon the location of the restriction site in the PCR product, which varies across populations. Therefore, measurement of these fragment lengths using electrophoresis provides information on microbial community diversity and

population dynamics. Another valuable component of RFLP and T-RFLP is that the PCR primer can be chosen to target a specific gene. Targeting a functional gene of interest allows monitoring of microbial populations capable of a specific function. In the denitrification pathway (Figure 1), primers have been developed for many of the steps, including *narG* (Philippot et al. 2002), *nirS/nirK* (Braker et al. 2000), *norB* (Braker and Tiedje 2003), and *nosZ* (Rosch et al. 2002; Rich et al. 2003).

The difference between RFLP and T-RFLP is in the number of nucleic acid fragment lengths that are measured for each population. In RFLP, every fragment is measured. On the other hand, T-RFLP utilizes a fluorescent-labeled primer that enables the identification of a single fragment for each microbial population. This allows for analysis of more complex communities and prevents double counting of populations that can occur with RFLP analysis. Both RFLP and T-RFLP can be visualized and quantified using electrophoresis, resulting in community composition data only including populations possessing the targeted gene. In the following section, I will discuss the analysis of fragment length data from these two techniques.

### 2.3.2 Electrophoresis Fragment Length Analysis

RFLP and T-RFLP provide a 'DNA fingerprint' of the microbial community that is visualized as an electropherogram (Figure 2). This electropherogram illustrates the separation of nucleic acid fragments of varying lengths via electrophoresis.

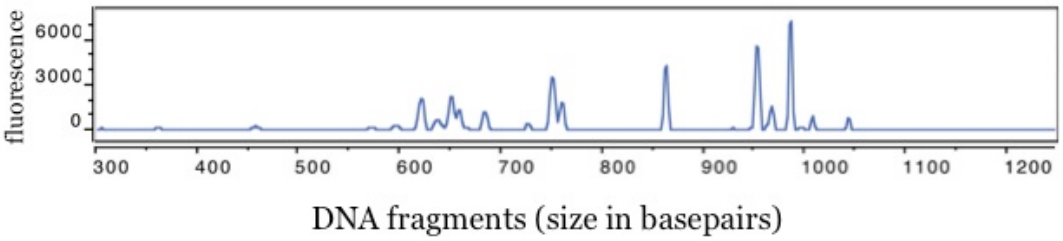



Figure 2. Electropherogram depicting different nucleic acid fragments present in a sample separated by electrophoresis. The area under each peak corresponds to the abundance of that nucleic acid fragment in the sample.

The locations of the peaks in the profile indicate nucleic acid fragments of various lengths. These fragments are each used as an identifier of a microbial population, also called an 'operational taxonomical unit' (OTU). The area under each peak correlates to the abundance of the OTU represented by that peak. Normalizing the peak areas by the sum of all peak areas yields a relative abundance for each OTU. This is done to allow comparisons between different samples, as the amount of product that is loaded into the capillary or gel varies and causes the total peak areas to vary.

In this thesis, electropherograms were created using both capillary and gel electrophoresis. Fragment lengths for the denitrifying microbial community analysis were measured using a capillary electrophoresis instrument. This instrument provides an electropherogram as seen in Figure 2. Gel electrophoresis was utilized to analyze microbial populations in the data obtained from Professor Hector Ayala-del-Rio (Ayala-del-Rio et al. 2004). In this technique, nucleic acid fragments of varying lengths appear as bands in a gel. These gels can then be scanned and converted into electropherograms using gel image analysis software.

## 2.4 Ecosystem Stability

Stability in ecology is a concept that does not have a clear definition, as many different parameters have been introduced to describe ecosystem stability (Grimm et al. 1992). These generally fall into two categories: stability of ecosystem functional performance and microbial community composition (Fernandez et al. 1999). Since engineered ecosystems have a designed function, the system performance provides a metric that can be used as a measure of ecosystem stability. This performance stability is of great concern to system operators and designers. Additionally, ecosystem stability can be studied by monitoring microbial population dynamics, whether the ecosystem has a measurable performance metric or not. In this section I will discuss stability concepts that have been used to analyze ecosystems, along with studies demonstrating the current understanding of functional and microbial community stability in microbial ecosystems.

### 2.4.1 Ecosystem Stability Measures

Of the many community composition stability concepts that have been used, Grimm and colleagues put forth the following four as the most important: constancy, resistance, resilience, and persistence (1992).

Constancy refers to a community remaining unchanged over time. Variability, the opposite of constancy, is more often used to quantify ecosystem stability, using measures such as the standard deviation and the annual variability of population abundances (Wolda 1978).

Resistance and resilience are concepts that refer to the response of an ecosystem to a disturbance. A measure of resistance, such as buffer capacity, quantifies the ability of a community to remain unchanged despite a disturbance (Jørgensen and Mejer 1977). On the

other hand, measures of resilience quantify the ability of an ecosystem to recover after a disturbance. One resilience measure is the domain of attraction, defined as the degree of disturbance that does not allow recovery (Grimm et al. 1992). Another commonly used measure of resilience is elasticity, which quantifies the length of time required for recovery after a disturbance (Grimm et al. 1992).

Persistence has been used in ecology as a measure of species longevity and can be quantified as mean time to extinction (Nisbet and Gurney 1982). Since extinction is not readily observed in long-lived species, the goals of persistence studies are often to model and predict extinctions based on population dynamics and factors influencing population fluctuations (Leigh 1981; Wissel and Stocker 1991).

### 2.4.2 Functional performance and microbial community stability studies

As engineered ecosystems rely heavily on microbial activity to achieve their designed function, it follows that the relationship between microbial community stability and functional performance stability would be of interest for analyzing and improving system design. Although many studies have examined microbial community structure and function, the nature of this relationship is still uncertain. Several studies of community dynamics in engineered ecosystems have shown that functional stability occurred without community composition stability (Fernandez et al. 1999; Smith et al. 2003; Cytryn et al. 2005; Gentile et al. 2006; Miura et al. 2007; Wang et al. 2011). The functional stability demonstrated in these studies include both general functions such as carbon removal (Fernandez et al. 1999; Smith et al. 2003; Miura et al. 2007; Wang et al. 2011) and more specific functions such as denitrification (Gentile et al. 2006), sulfide oxidation (Cytryn et al. 2005), and total nitrogen removal (Wang et al. 2011). This functional stability suggests a

level of redundancy and interchangeability of the microbial populations in performing the system function. While it has been suggested that microbial diversity leads to more stable systems (Daims et al. 2001; Rowan et al. 2003; Girvan et al. 2005), stability has also been linked to specific populations in a microbial community rather than diversity (Griffiths et al. 2004). This suggests that the composition of the microbial community is more important than the number of populations to the stability of the system. The idea that less diversity can lead to more stable performance has been demonstrated in reactors operated under different conditions to achieve varying levels of microbial diversity (Pholchan et al. 2010). Other studies have shown that unstable functional performance can occur both with stable and unstable community composition (Martienssen and Schops 1997; von Canstein et al. 2001; Stamper et al. 2003).

Clearly, there is no consensus as to the nature of the relationship between microbial diversity and ecosystem stability, or even as to what constitutes a stable ecosystem. With these uncertainties in mind, we propose the use of reliability theory as an alternative approach that links ecosystem functional stability with microbial community stability and diversity.

## 2.5 Reliability Theory

Reliability Theory is a branch of statistics and probability that deals with failures of systems and components. Although the foundations of the discipline of reliability engineering were set by the 1930s, it did not emerge as a major area of study until the 1950s. The driving force behind this new discipline was the post-World War II desire for greater quality control of manufactured products, especially electronic components (Saleh

and Marais 2006). The tools used in reliability theory and the possible applications have evolved greatly over time. Reliability theory is now not only widely used in electronics production, but can also be seen in military equipment production (Neil et al. 2001; Adamides et al. 2004), computer science (Moranda 1975; Goseva-Popstojanova and Trivedi 2000), medicine (Green and Hedinger 1978; Gavrilov and Gavrilova 2001;; Tees et al. 2001; Korell et al. 2011), biological sciences (Gavrilov and Gavrilova 2001; Laird and Sherratt 2009; Laird and Sherratt 2010), and structural engineering (Rackwitz and Fiessler 1978; Enevoldsen and Sorensen 1994; Vu and Stewart 2000). In this study we have developed a novel approach to apply reliability theory to an engineered ecosystem to shed light on potential improvements in system design.

Reliability theory uses statistical methods to predict failure events. To make these predictions, identical components are tested under typical operating conditions until failure in what is called a life test. The length of time to a failure event for a test subject is often called a 'lifetime', 'time to event', or 'time to failure' in reliability theory.

In the following sections, I will describe the steps in modeling the reliability of a system. First I will define the many functions that are used in reliability theory. I will then introduce some common failure distributions that are used to model component failures. Next I will discuss how the parameters in these failure distributions are estimated from life test data. I will then introduce different system component configurations and the reliability calculation of each. Finally, once a system reliability model is complete, I will discuss how to analyze the model to determine the reliability-limiting component.

### 2.5.1 Reliability Functions

There are many functions used in reliability theory that require definition. Reliability, $R(t)$, defined as the probability that a component will operate properly for time *t*, can be calculated as the ratio of surviving components to total components at time *t*.

$$R(t) = \frac{\#\text{ of surviving components at time } t}{\#\text{ of total components}} \quad (1)$$

If $T$ is a random variable denoting the 'time to event', this reliability function can also be described as the probability that $T$ is greater than time *t*. Often this is also called the Survivor Function, $S(t)$.

$$S(t) = R(t) = P(T > t) \quad (2)$$

The cumulative distribution function of failure, $F(t)$, is the complement of the reliability and gives the probability that $T$ is less than time *t*.

$$F(t) = 1 - R(t) \quad (3)$$

The probability density function or event density, $f(t)$, is then defined as the derivative of the cumulative distribution function. This gives the probability that a component will fail at exactly time $t$.

$$f(t) = \frac{dF(t)}{dt} \quad (4)$$

The hazard function or instantaneous failure rate, $h(t)$, is defined as the probability density function divided by the reliability:

$$h(t) = \frac{f(t)}{R(t)} \qquad (5)$$

The hazard function is the conditional probability of failure in the interval $t$ to $t + \Delta t$, given that there was no failure at time *t*. This is also called the 'instantaneous failure rate function' and is useful in estimating time to failure, among other things.

### 2.5.2 Failure Distributions

As the application of reliability theory has expanded, so have the failure behaviors that these applications exhibit. Because of this, many different failure distributions have been developed to model and predict system failures. In the current analysis, measured component times to failure are continuous variables and therefore continuous failure distributions were employed rather than discrete (Table 1). The reliability analysis software tool developed in this thesis allows for any of the failure distributions in Table 1 to be fit to life test data.

Table 1. Continuous failure distribution models commonly used in reliability engineering and their characteristics.

| Model | $f(t)$ | Characteristics | Applications |
|---|---|---|---|
| Weibull | $\frac{\gamma}{\theta}\left(\frac{t}{\theta}\right)^{\gamma-1}\exp\left[-\left(\frac{t}{\theta}\right)^{\gamma}\right]$ | Can have increasing, constant, or decreasing failure depending on value of gamma. | Very versatile; Originally designed for particle size distribution (Rosin and Rammler 1933) and material strength (Weibull 1939). |
| Rayleigh | $\frac{t}{b^2}e^{\left(\frac{-t^2}{2b^2}\right)}$ | Linearly increasing hazard function | Used for systems described by intense aging effects |
| Lognormal | $\frac{1}{\sigma t\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\ln t-\mu}{\sigma}\right)^2\right]$ | Logarithms of data values are normally distributed | Often used in materials strength, fatigue and general reliability analysis |
| Extreme Value | $\sigma^{-1}\exp\left(\frac{t-\mu}{\sigma}\right)\exp\left(-\exp\left(\frac{t-\mu}{\sigma}\right)\right)$ | Highly increasing failure rate | Components showing significant wear and tear. |
| Exponential | $\frac{1}{\mu}e^{-\frac{t}{\mu}}$ | Constant failure rate | Components without a mechanism for wear and tear. |
| Normal | $\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(t-\mu)^2}{2\sigma^2}}$ | Random variables clustered around a mean value | Very commonly used distribution |

### 2.5.3 Parameter Estimation Methods

Once a failure distribution model in chosen, the parameters in the model must be determined using the results from a life test. As discussed earlier, a life test results in 'time to failure' data for identical units tested under typical operating conditions. This failure data can be of two types: complete or censored. When failures were observed and quantified for all test subjects, it is referred to as complete data. If only partial information about the failure times of a test subject are known, for example when a life test is terminated without a test subject reaching failure (Figure 3), then the data is said to be censored. This type of censoring is called right censoring. The failure time of this subject is not known, although it is known to be greater than the time at which the test was terminated.

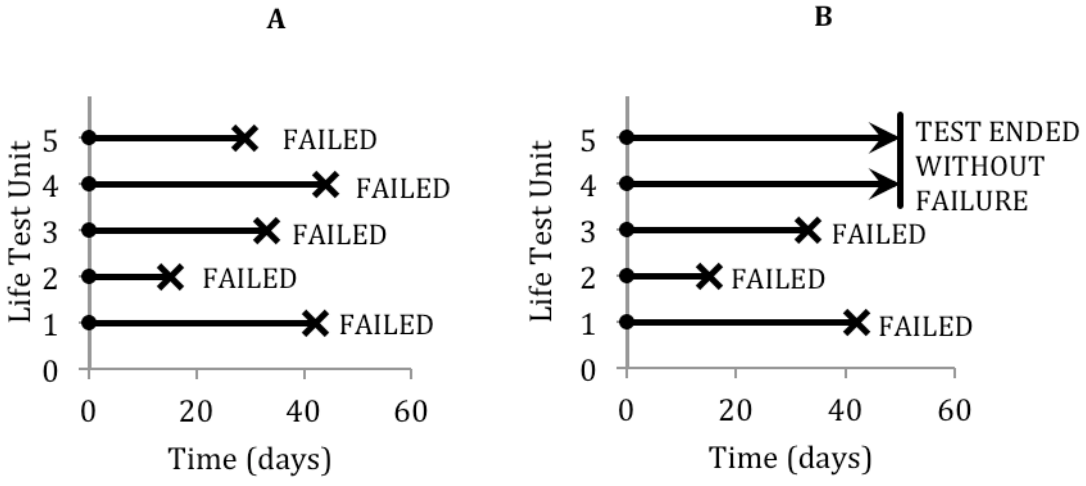



Figure 3. Graphical depiction of life test results consisting of complete and right-censored data. If failures are observed for all units tested (A), the failure data is said to be complete. Otherwise, if a life test is terminated before some of the test units reach failure (B), the data is said to be right-censored.

Whether life test data is complete or censored must be taken into account when estimating reliability distribution parameters to obtain an accurate reliability function. There are several parameter estimation methods that have been developed, including the method of moments, method of least squares, and maximum likelihood estimation, and not all of them handle censored data well.

The method of moments is a method of parameter estimation that equates moments of statistical characteristics of a sample to moments of the same characteristics of a population distribution. Solving these equations then yields the estimated population distribution parameter values. This method is mathematically simple and can be done without extensive computations, but it does not have the desired asymptotic properties of other methods and is best applied to complete data. Because of this, the method of moments has been largely replaced with more robust parameter estimation techniques.

The method of least squares is a parameter estimation method that linearizes the reliability model distribution and uses linear regression to find the distribution parameters that result in the best fit. Like the method of moments, the method of least squares is best used with complete data. It can also only be used with distributions that can be linearized, which prevents its use in some models.

In this analysis, maximum likelihood estimation is used, as it is the most robust of the parameter estimation techniques. Although maximum likelihood estimation can handle censored data, there are still limits to the amount of censoring that can be present in the dataset. In this thesis, at least one 'time to failure' must not be censored and there must be two total 'times to failure' for parameter estimation. The maximum likelihood estimation technique used here is described in detail in the following section.

#### 2.5.3.1 Maximum Likelihood Estimation

Maximum likelihood estimation (MLE) is one of the most powerful tools for fitting reliability distributions to data. This method allows the estimation of distribution parameters, and in turn, values for mean and variance, using only a subset of a population.

MLE begins with the creation of a likelihood function from an observed data sample and a failure distribution model such as those in Table 1. This creates a unique likelihood function that relates parameter values in the failure distribution to the probability these parameters would give rise to the observed data.

For a sample containing right-censored data, the data can be split into two disjoint sets: U with uncensored data and C with right-censored data. The likelihood function for a two-parameter distribution would then be:

$$L(\theta,\gamma) = \left\{ \prod_{i \in U} f(t_i,\theta,\gamma) \right\} \left\{ \prod_{i \in C} S(t_i,\theta,\gamma) \right\} \quad (6)$$

where $f(t)$ is the probability density function, $S(t)$ is the survivor function, and $\theta$, $\gamma$ are distribution parameters.

Maximum likelihood estimation attempts to make the best estimate of the parameters in the failure distribution using the likelihood function. As discussed earlier, the likelihood function relates parameter values in the failure distribution to the probability that these parameters would give rise to the observed data sample. Therefore, the parameter values that maximize the likelihood function are the values most likely to

result in the observed data. In order to maximize the likelihood function, most often the logarithmic expression of the likelihood function is first taken, as this transforms the task of maximizing the likelihood function from maximizing a product to maximizing a sum, which simplifies it greatly.

$$l(\theta,\gamma) = \log L(\theta,\gamma) \tag{7}$$

This log-likelihood function is then differentiated with respect to its parameters. The differential equations are set equal to zero and solved simultaneously to obtain the estimates of the parameter values that maximize the likelihood function.

$$\frac{\delta l}{\delta \theta_i} = 0 \qquad i = 1,2,\ldots,m \tag{8}$$

Although solving this set of differential equations might not be trivial, it can be accomplished with adequate computational ability, resulting in the best estimates for the failure distribution model parameters.

#### 2.5.3.2 Goodness-of-Fit Tests

As seen in Table 1, there are many different failure distribution models to choose from when modeling reliability data. Therefore, after using life test data to determine model parameters with maximum likelihood estimation or another parameter estimation technique, the goodness-of-fit of the chosen distribution must be analyzed to determine if the model accurately describes the failure data. There are several goodness-of-fit tests that have been developed to indicate whether a sample comes from a population with a given

distribution, including the Kolmogorov-Smirnov (K-S) test, the Anderson-Darling test, and the chi-square test.

The Kolmongorov-Smirnov test is a powerful goodness-of-fit test that is based on the empirical distribution function and can be used with even small sample sizes (Chakravarti et al. 1967). One major limitation of the K-S test is that the population distribution must be known, meaning that if the distribution parameters were estimated using the failure data, the test is no longer valid. In this analysis the parameters are estimated from the failure data and therefore the K-S test could not be used to determine goodness-of-fit.

The Anderson-Darling test is a modification of the K-S test (Anderson and Darling 1952). It compares a test statistic to a critical value in order to determine whether a sample is from a specified distribution. Unlike the K-S test, this test can be used with estimated parameters, but the critical values that the test statistic is compared to must be calculated for each different distribution. These values are tabulated and can be found in the literature (Stephens 1974; Stephens 1976; Stephens 1977a; Stephens 1977b; Stephens 1979;). The test statistic for the Anderson-Darling test is calculated as:

$$A^2 = -N - S \tag{9}$$

$$S = \sum_{i=1}^{N} \frac{(2i-1)}{N} [\ln F(Y_i) + \ln(1 - F(Y_{N+1-i}))] \tag{10}$$

where $F$ is the cumulative distribution function, $N$ is the sample size, and $Y_i$ is the ordered data.

While the Anderson-Darling test has the advantage of being applicable when the failure distribution parameters were estimated from the data whose fit is being analyzed, its use can become cumbersome when several different models are being tested, as both the test statistic equation and the critical value tables must be modified for each distribution type used. For this reason, the Anderson-Darling test was not used in this thesis.

The chi-square test is a goodness-of-fit test that uses binned data and can be applied to both continuous and discrete distributions (Snedecor and Cochran 1989). The chi-square test compares a test statistic to the chi-square percent point function to test the null hypothesis that the data follow the specified distribution versus the alternate hypothesis that it does not. The test statistic is calculated as:

$$\chi^2 = \sum_{i=1}^{k} \left(O_i - E_i\right)^2 / E_i \qquad (11)$$

where $k$ is the bin number, $E_i$ is the expected frequency for bin *I*, and $O_i$ is the observed frequency for bin *i*.

The expected frequency, $E_i$, for each bin is calculated as:

$$E_i = N(F(Y_u) - F(Y_l)) \qquad (12)$$

where $F$ is the cumulative distribution function, $Y_u$ is the upper limit for bin *i*, $Y_l$ is the lower limit for bin *i,* and $N$ is the sample size.

The null hypothesis is then rejected if

$$\chi^2 > \chi^2_{(\alpha,k-c)} \quad (13)$$

where $\alpha$ is the significance level, $k$ is the number of non-empty cells, $c$ is the number of estimated parameters plus one, and $\chi^2_{(\alpha,k-c)}$ is the chi-square percentage point function for given variables.

Like the Anderson-Darling test, the chi-square test has the advantage of being applicable when the failure distribution parameters were estimated from the data whose fit is being analyzed. A further advantage to the chi-square test is that its test statistic equation and critical values are not dependent on the failure distribution model being analyzed. This allows for easier integration into the reliability analysis software tools developed in this thesis. The chi-square test also has some disadvantages; including the effect that different binning schemes have on the results and that a sufficiently large sample size is required for the test to be valid. In the interest of software development, the chi-square test was implemented in the reliability analysis methods in this thesis.

### 2.5.4 Components and Systems

In engineered production systems, two of the most important design considerations are efficiency and reliability. Production at the fastest rate or lowest cost while minimizing downtime is often the goal in system design. The desire for efficiency has given rise to specialization of components, creating systems of components in series. On the other hand, minimizing system downtime during inevitable component failures has led to the use of parallel components of redundant function. Therefore, often the optimal system design will

consist of both series and parallel components. A similar situation can be seen in ecological engineering. Clearly specialization of function occurs among microorganisms. For example, within the denitrification pathway discussed earlier, some microbial populations may only be capable of one step in the reduction. The denitrification would then be completed by some other microorganisms that possess the necessary functional genes. The functional diversity found at each step can also be viewed as parallel components in the microbial 'system'. This approach to modeling engineered ecosystems as microbial populations in series and parallel will be discussed further in Section 3.1.1.

To examine complex component systems with an infinite number of possible component configurations, first we must look at the simplest of configurations: series systems and parallel systems.

#### 2.5.4.1 Series Systems

Series systems make up the simplest multi-step systems. In series systems the output from one component is the input for the following component (Figure 4).

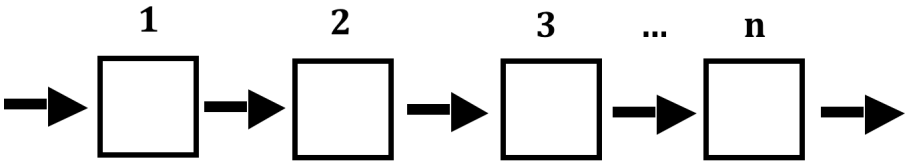



Figure 4. Series system block diagram. In series systems, the output from one component is the input for the following component.

In the case of a simple series system, the reliability of the system is equal to the probability that all of the components are operational. This is expressed as:

$$R = \prod_{i=1}^{n} P(x_i) \qquad (14)$$

where $P(x_i)$ is the probability of the *i*th unit being operational.

#### 2.5.4.2 Parallel Systems

Simple parallel systems consist of one set of parallel components (Figure 5). The parallel components in these systems are redundant in function, meaning that only one of such components must be functioning for the system to function.

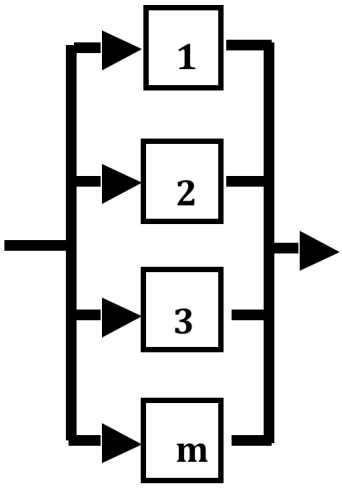



Figure 5. Simple parallel system block diagram. In a simple parallel system, components in parallel are of redundant function.

For a parallel system to fail, all of the components must fail. Therefore, the reliability of a parallel system is the complement of the probability that all of the parallel components have failed. This is expressed as:

$$R = 1 - \prod_{i=1}^{n} P(\overline{x_i}) \qquad (15)$$

where $P(\overline{x_i})$ is the probability of failure of the *i*th unit.

#### 2.5.4.3 Complex Systems

More complex component systems arise from the combination of parallel and series components (Figure 6). In calculating the reliability of this system, first the reliability of each subsystem of *m* parallel components is calculated just as in the simple parallel system, and then these reliability values are used for each *n* subsystem in the simple series system equation. The same approach of breaking down complex systems into subsystems to calculate the overall system reliability can be used on even more complex systems.

.

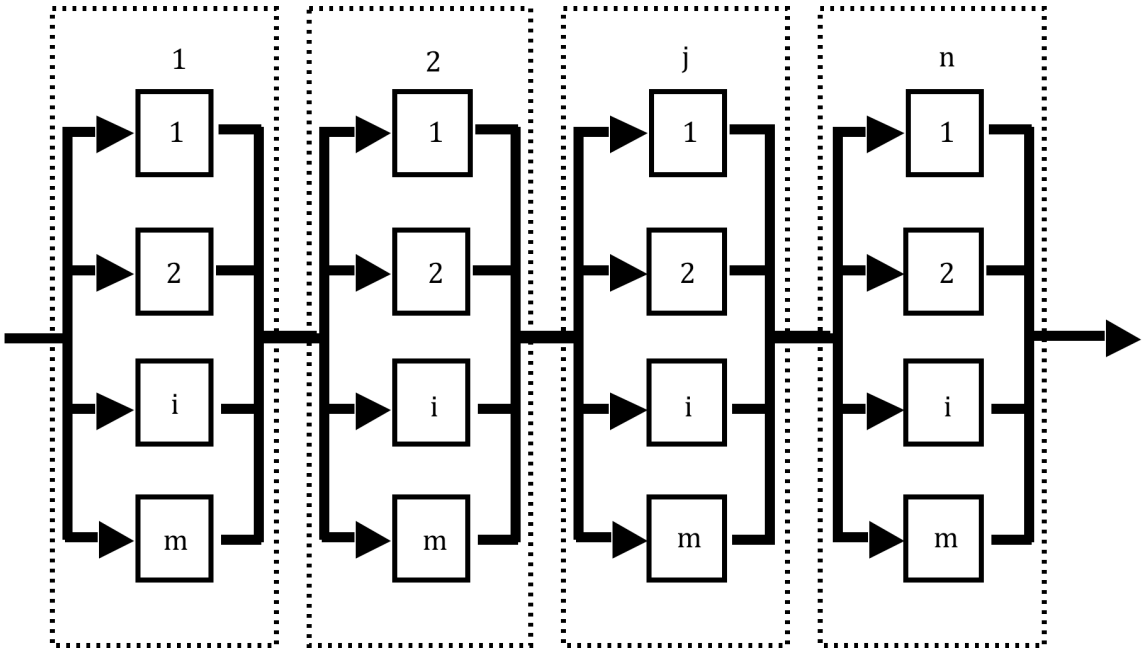


Figure 6. Series-parallel system block diagram. Series-parallel systems have several steps and redundancy at each step.

The reliability for this system is expressed as:

$$R = \prod_{i=1}^{n}\left[1 - \prod_{j=1}^{m}(1 - P(x_{ij}))\right] \qquad (16)$$

where $P(x_{ij})$ is the probability that unit *ij* is operational.

Modeling an ecosystem as a system of components where each component is a microbial population can give rise to complex systems like those seen in Figure 7, as many microbes are capable of multiple functions in a system while others are only capable of one. This may lead to some components spanning several steps in the functional series and some spanning as few as one. Although the system configurations can become quite complex, as long as they can be broken down into simple series and parallel components, Equations 14 and 15 can be used and analysis of the system is fairly straightforward. There

are instances, however, where these equations are not adequate and system analysis becomes more complicated. One such occurrence is *k*-out-of-*n* systems, which are described in the following section.

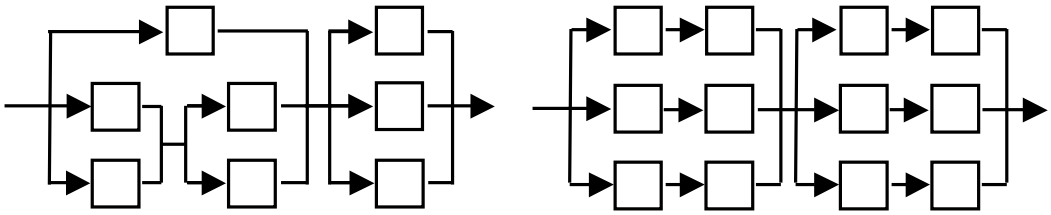


Figure 7. Complex component systems with varying configurations of parallel and series components.

#### 2.5.4.4 *k*-out-of-*n* Systems

In the parallel system described earlier, one only parallel component must be operational for the system to function; therefore failure occurs when all of the parallel components fail simultaneously. In many real-life systems, successful system function requires more than one parallel component to be operational. Examples of this include a V8 automobile engine that requires at least 4 working cylinders or a jet plane that requires 2 of its 4 engines. In both of these examples, the systems can tolerate some component failures. Systems of this type are called *k*-out-of-*n* systems, where *k* is the number of components that are required to be operational out of the *n* total components.

If the components of the system have identical and independent life distributions, determining the reliability is fairly straightforward; it is the probability that at least *k* units are functioning:

$$R(k;\, n, p) = \sum_{r=k}^{n} \binom{n}{r} p^r (1-p)^{n-r} \qquad (17)$$

where *p* is the probability that a unit is operational.

If the components of the *k*-out-of-*n* system are not identical, determining the system reliability is more complex and requires greater computational power. Several algorithms have been developed which make this task more efficient (Barlow and Heidtmann 1984; Rushdi 1986; Belfore 1995). Barlow and Heidtmann (1984) developed the following recurrence relation that can be used to determine the reliability of *k*-out-of-*n* systems where components have different reliabilities:

$$R_e(i,j) = q_j R_e(i,j-1) + p_j R_e(i-1,j-1), \quad 0 \le i \le n, 0 \le j \le n \tag{18}$$

With boundary conditions:

$$R_e(-1,j) = R_e(j+1,j) = 0 \quad \text{for } j = 0,1,2,... \tag{19}$$

$$R_e(0,0) = 1 \tag{20}$$

where $R_e(i,j)$ is the probability that exactly $i$ out of $j$ components of the system are operational, $p_j$ is the reliability of component $j$, and $q_j$ is the unreliability of component $j$.

The system reliability can then be determined using the following equation:

$$R(k,n) = \sum_{i=k}^{n} R_e(i,n) \tag{21}$$

An even more efficient use of this method can be employed if it is only of interest to determine the probability that at least $k$ units are operational. In this case, calculating $R_e(i,j)$ can be avoided and the reliability of the system can be determined using:

$$R(k,k) = R_e(k,k) \tag{22}$$

$$R(k,j) = R(k,j-1) + p_j R_e(k-1,j-1) \quad \text{for } k > j \tag{23}$$

The recurrence relation in Equation 18 results in a value for both the probability that at least *k* out of *n* components are operational, which is the system reliability, along with all of the intermediate $R_e(i,j)$ values. These intermediate values can also be useful in analyzing system design, and therefore the added complexity of Equation 18 can be justified in some instances.

### 2.5.5 Reliability Importance of Components

The goal of modeling systems using reliability theory is often to determine avenues of improving system design and operation. Therefore, determining which system component is limiting the overall system performance is of interest. The concept of the reliability importance of components has been developed to indicate which system component is limiting. The reliability importance of component *k*, $I_k(t)$, is calculated by taking the partial derivative of the system reliability by the reliability of component *k*. Therefore, an improvement in the reliability of the component with the highest reliability importance would yield the largest improvement in overall system reliability.

$$I_k(t) = \frac{\partial R_S(t)}{\partial R_k(t)} \tag{24}$$

where $\partial R_S(t)$ is the system reliability at time *t* and $\partial R_k(t)$ is the reliability of component *k* at time *t*.

In the reliability software tool developed in this thesis, the reliability importance can be analyzed for groups of components to determine which is limiting the system reliability. Applying this to the denitrifying biofilters can provide information on the rate-

limiting step in the denitrification pathway. If the limiting step can be identified, this step can then be targeted to determine avenues of improving the reliability of the system.

# Chapter 3: Ecosystem Reliability Analysis

In this section, I will first discuss the method used to apply the reliability concepts introduced earlier to engineered ecosystems (Figure 8). I will then present the Ecosystem Reliability Analysis Tool (EcoReliAnT), a toolbox that I created in MATLAB® to allow users to easily analyze microbial community data using the reliability method presented in this thesis. The following two sections will then demonstrate the use of the reliability analysis and EcoReliAnT on two different datasets: the first being from a published study and the second from field denitrifying biofilters.

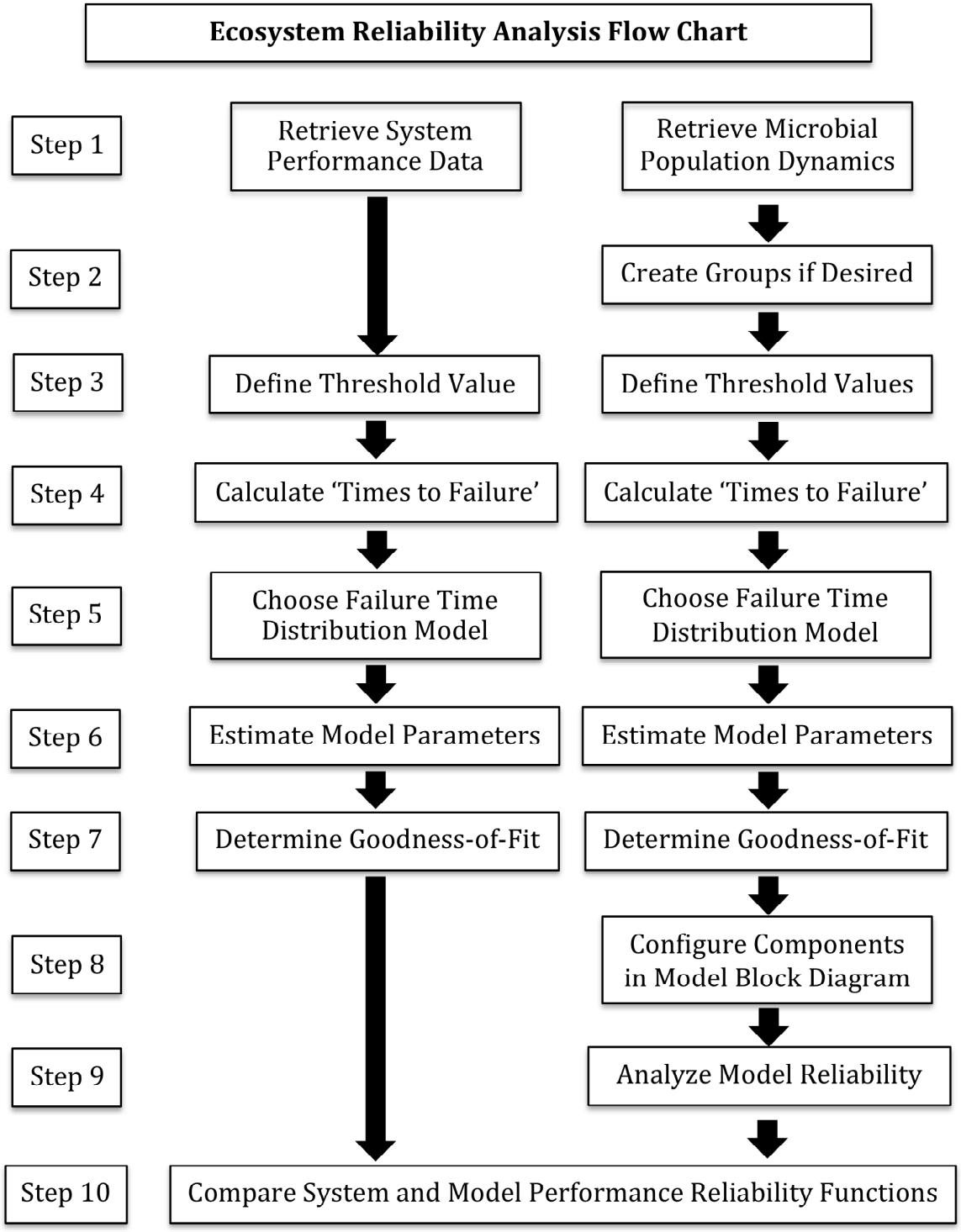



Figure 8. Ecosystem reliability analysis method flow chart.

## 3.1 Methodology

### 3.1.1 Microbial Communities as a System of Components

Microbial populations were represented as components in a system to facilitate the application of reliability theory to engineered ecosystems. Using this approach, microbial populations involved in performing the system function were arranged in a reliability block diagram. The diversity of microbial populations capable of performing the system function was modeled as parallel components in the system. In the two analyses in this thesis, the available microbial community information resulted in systems with only one step in their function. However, in systems whose function involves several steps, each step could be arranged in series with the populations in subsequent steps. For example, in our proposed reliability block diagram of the denitrifying biofilers, each step in the reduction of nitrate to dinitrogen gas is performed by groups of components in series, with fungal activity modeled as a peripheral role in the conversion of woodchips into readily available carbon compounds for the microbes (Figure 9). However, to date only the *nosZ* microbial community data was available and therefore the denitrifying biofilter system was modeled based on the final reduction step. The diversity at this step was acquired using the T-RFLP microbial fingerprinting technique (Liu et al. 1997). Functionally redundant microbial populations that were correlated to system performance were considered to be parallel components at this step. Each of the parallel components consisted of an individual microbial population described by a unique reliability function. I have developed a method to quantify these microbial reliability functions based on the population dynamics information obtained using microbial fingerprinting techniques.

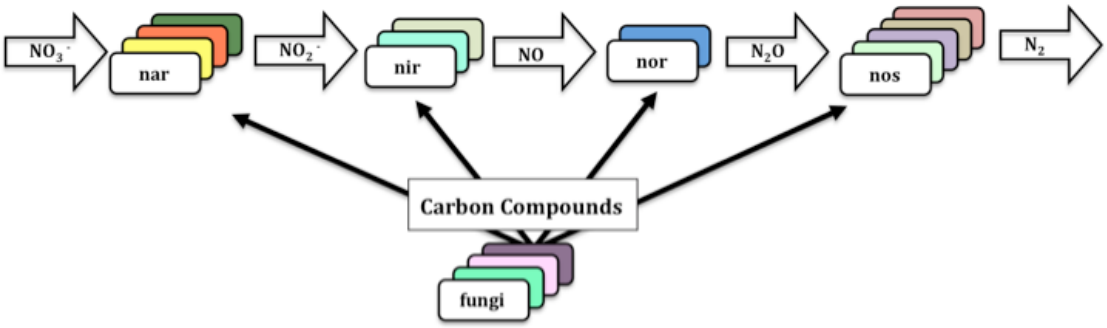



Figure 9. Proposed denitrifying biofilter reliability block diagram with the different genes responsible for each step in the denitrification process modeled as groups of components in series. The number of microbial populations possessing each of these genes corresponds to the number of parallel components in each group.

### 3.1.2 Determining Microbial Reliability Using Population Dynamics

To analyze the dynamics of a microbial population, its abundance was monitored over time using microbial fingerprinting techniques (Figure 10). This yields a relative abundance versus time for any given microbial population (Figure 11). As discussed earlier, reliability functions are determined using the results of life tests where components are operated until failure, with the time elapsed until this failure called a 'lifetime' or time to failure for that component. Because in biology "lifetime" is commonly used at the organism rather than the population level, which could lead to confusion, we use the "time to failure" term here. In our reliability analyses, we assumed the length of time a microbial population is continuously present in an RFLP or T-RFLP profile to be a time to failure for that population. Multiple times to failure for a population comprise the failure data utilized for maximum likelihood estimation of parameters describing a population.

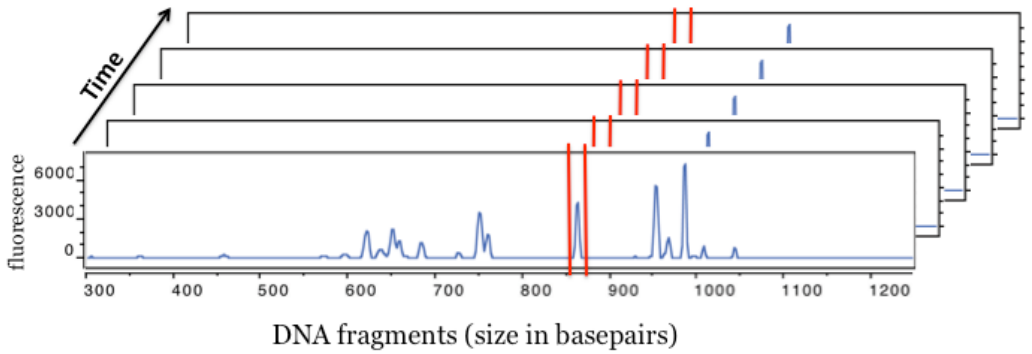



Figure 10. Monitoring population dynamics across electropherograms from temporal samples provides information about temporal changes in a microbial population's relative abundance in the system.

In quantifying these times to failure, we have also implemented the use of an adjustable threshold to define a relative abundance value that is required for a population to be considered present in a sample (Figure 11). This allows for the possibility that at a low abundance, a population's impact on system performance may be negligible, although that population is never be completely absent from the community. Therefore, at relative abundances below the threshold, the system component represented by this microbial population has 'failed' in its role in the system function. Using this type of threshold, the length of time from when a microbial population's relative abundance surpasses the defined threshold until it drops below the threshold was considered one time to failure. If the same population subsequently surpasses the threshold again, a second life test would result.

Some microbial populations may arise in an ecosystem that cause poor performance. These types of populations can be thought of as 'nuisance' populations. Whether they are hindering system performance directly or out-competing the desired populations, they can affect the reliability of the system. In our reliability analysis, if a

population is determined to be negatively correlated with system performance, the spans of absence above the threshold are used as 'times to failure' rather than the spans of presence. This means that the reliability function determined from the nuisance population's time to failure data is actually the probability that the nuisance population is absent in the system, and failure occurs when the nuisance population is present. This approach is used throughout the thesis to handle microbial populations that are negatively correlated with system performance.

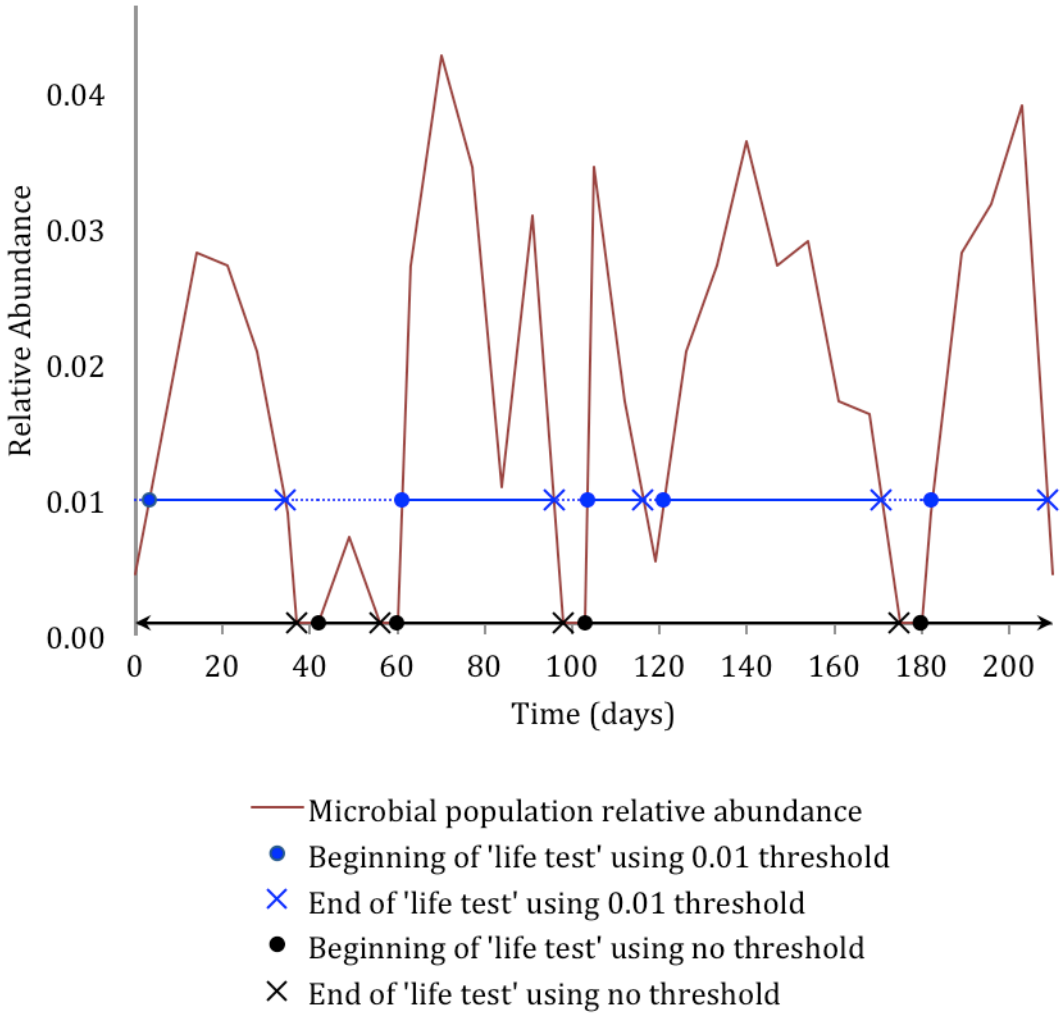



Figure 11. 'Times to failure' for microbial populations are determined from population dynamics. For a population that is positively correlated to system performance, spans of presence above a defined threshold are seen as individual 'life tests' each yielding a 'time to failure'. For populations negatively correlated with system performance, 'time to failure' data is determined from spans of absence, as the presence of these 'nuisance' populations are seen as causing system failure. The 'times to failure' for these populations are indicated by the dashed lines.

After the 'time to failure' dataset has been obtained from the dynamics of a microbial population, this data is used to estimate the parameters of a failure distribution model for the corresponding microbial population using maximum likelihood estimation. The shapes of failure distribution model curves can vary greatly, even when parameters are estimated

from the same data (Figure 12). Therefore, after using life test data to determine model parameters with maximum likelihood estimation or another parameter estimation technique, the goodness-of-fit of the distribution must be analyzed to determine if the model accurately describes the failure data. In the software tools developed in this thesis, the goodness-of-fit of reliability functions are automatically assessed using the chi-square test. Along with the test statistic, the chi-square test provides a p-value that allows for comparison of the goodness-of-fit of different models to a given dataset. However, since the sample sizes analyzed in this thesis were small, the chi-square test was not valid and therefore the Weibull model was used for all reliability functions because of its versatility.

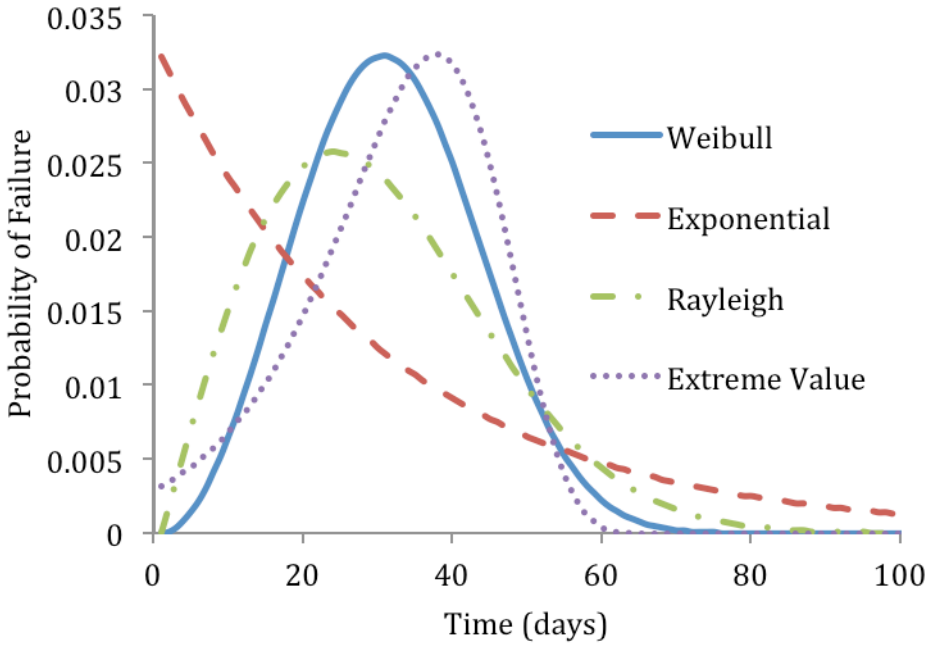



Figure 12. Different probability density function models with parameters estimated from the presence above the 0.01 threshold time to failure data in Figure 10. Using a goodness-of-fit test can determine whether a model accurately represents the failure data it is estimated from.

Since the reliability methods presented in this thesis are computationally intensive and time consuming when applied to large datasets such as ours, software tools are necessary to implement them. The following section describes the development of a software tool to apply these reliability methods to ecosystem data.

## 3.2 Development of the Ecosystem Reliability Analysis Tool

Since this is a new approach to quantify ecosystem stability, tools were not previously available to implement the analysis method. Software in MATLAB® (Version R2009a, The MathWorks, Inc., Natick, MA) was created that allows us to rapidly analyze microbial population dynamics, determine microbial reliability functions, and analyze system component configurations by implementing the analysis steps detailed in Figure 8. This MATLAB® software utilizes the Statistics Toolbox™ and the GUI Design Environment (GUIDE) tool in MATLAB® to create a graphical user interface that incorporates all of the code used in this analysis (The MathWorks, Inc., Natick, MA). This section describes the development of the Ecosystem Reliability Analysis Tool (EcoReliAnT) and how it was used to analyze microbial ecosystems in this thesis. For further information on the EcoReliAnT interface, see the Appendix. The development of EcoReliAnT consists of several major functions: (1) Importing data into the tool; (2) Filtering and grouping microbial populations; (3) Creating presence threshold for the determination of 'times to failure'; (4) Determining reliability functions for microbial populations based on 'time to failure' data; and (5) Configuring microbial components in a system and analyzing the fit of this model to the system performance function.

### 3.2.1 Importing Data

The first step in the development of EcoReliAnT was to implement code that allows microbial fingerprinting data to be imported from the MATLAB® workspace. This data consisted of two parts: a numerical array and a cell array of strings (Figure 13). In the numerical array, the first column consisted of the numbered sample days while subsequent columns contained the time series abundance data, with each column including data for a different microbial population. Each row in this array corresponded to a single sampling date. The cell array of strings consisted of the headers for each of the columns in the numerical array, with each cell number in the cell array matched to its corresponding column in the numerical array. EcoReliAnT allows the user to specify the names of these two arrays to import their data for further analysis.

| CELL ARRAY OF STRINGS CONTAINING COLUMN HEADERS | | | | | |
|---|---|---|---|---|---|
| Sample Days | Population 1 Name | Population 2 Name | Population 3 Name | Population 4 Name | Population 5 Name |
| **NUMERICAL ARRAY CONTAINING DAY NUMBERS AND ABUNDANCE DATA** | | | | | |
| N | P | P | P | P | P |
| U | O | O | O | O | O |
| M | P | P | P | P | P |
| B | | | | | |
| E | 1 | 2 | 3 | 4 | 5 |
| R | | | | | |
| E | A | A | A | A | A |
| D | B | B | B | B | B |
| | U | U | U | U | U |
| S | N | N | N | N | N |
| A | D | D | D | D | D |
| M | A | A | A | A | A |
| P | N | N | N | N | N |
| L | C | C | C | C | C |
| E | E | E | E | E | E |
| | | | | | |
| D | D | D | D | D | D |
| A | A | A | A | A | A |
| Y | T | T | T | T | T |
| S | A | A | A | A | A |

Figure 13. Data format required for import into EcoReliAnT. Input data for EcoReliAnT must be formatted to consist of two arrays: a numerical array containing the time series abundance data for each microbial and numbered sample days and a cell array of strings containing the column headers for each column in the numerical array.

### 3.2.2 Filtering and Grouping Populations

Since hundreds of microbial populations were present in our denitrifying biofilter samples and not all could be used to model the reliability of the system, I needed a way to filter the imported data to exclude unwanted populations from the analysis. This was achieved by allowing the user to define both a number of times a population must appear in the time series and an abundance value requirement that the population must reach to be included in the data to be analyzed further.

Another desired function was the ability to group populations together. This was utilized in my analysis to combine datasets from microbial populations with similar dynamics, allowing for more 'times to failure' to be observed. This larger sample size provides a more accurate reliability function for the population group. EcoReliAnT allows the user to input populations to be grouped together as text in a cell array to define the groups. These populations can then either be grouped by combining their abundance values or by combining the 'time to failure' datasets generated for the individual populations.

### 3.2.3 Creating Presence Thresholds and Calculating 'Times to Failure'

After importing and grouping the data, the next step in the analysis requires threshold values to be defined for determination of spans of presence and absence. Thresholds can be defined in three different ways. The first method of defining a threshold is to use one user-specified value for all populations or groups. The second method is to manually input different threshold values for each group or population. The third method is to create a threshold for each population or group based on the average and standard deviation of its abundance value. This method calculates thresholds based on an input value that is multiplied by the standard deviation and then added to the average value.

Once the threshold value is defined, the next step is to calculate the spans of presence and absence based on the population data and threshold value. Calculating the spans of presence and absence simultaneously in this step allows the user to choose which to use as 'times to failure' later in the analysis. As discussed earlier, using the spans of absence of a population as its 'time to failure' data is useful for 'nuisance' populations whose presence causes the system to fail. Two options are available for determining

abundance values between sample dates and therefore the point of intersection with the threshold: linear interpolation between points or a stepwise function.

### 3.2.4 Determining Reliability Functions for Populations

After calculating spans of presence and absence for a population or group, the maximum likelihood estimation (mle) function in MATLAB® is used to determine reliability function parameters and their 95% confidence intervals. At this point, the user can choose whether the 'times to failure' used to develop the reliability function consist of the spans of presence or spans of absence of the population or group of interest. Along with the failure data, EcoReliAnT requires two additional inputs to develop a reliability function. First, a reliability model must be chosen from the many options in the mle function. Second a censoring vector must be created for input in the mle function to identify which of the 'times to failure' are right-censored. EcoReliAnt automatically creates this censoring vector when calculating the 'times to failure' from the population data. In order to accommodate these inputs, EcoReliAnT allows the user to choose the reliability model from a list and choose whether or not to include the censoring vector in the analysis. Choosing not to include the censoring vector would consider all failure data to be complete data when determining the reliability function.

Once parameters for a reliability function have been estimated using the mle function, the goodness-of-fit of each model is automatically analyzed using the chi square goodness-of-fit (chi2gof) function.

After a suitable reliability function is determined from the failure data, this function needs to be saved for future use. A 'saved fit bank' was developed that stores reliability functions, allowing them to be retrieved later in the analysis.

### 3.2.5 Configuring Components in System and Analyzing Model Fit

Once the reliability functions for microbial populations or groups were determined, we needed a way to assign specific reliability functions to components in a system. We also needed a way to specify the configuration of each of these components in the block diagram to solve for the reliability of the system. For this purpose, the concept of 'structures' and 'nodes' was developed. A 'structure' is a set of components and can be made up of any components in the system that have input and output at the same location. A structure can be a single component, a group of components in parallel, a *k*-out-of-*n* system, or what is called a bridge. For any *k*-out-of-*n* systems used, the *k* value for each structure must also be defined. The input and output locations of the structures are specified by 'nodes'. Nodes are numbered markers that specify the configuration of components in the system. For example, if two structures both have their input at node 1 and output at node 2, EcoReliAnT can identify these as parallel structures. Bridges are structures that are used to define complex systems whose configurations cannot be identified using nodes alone. For a more detailed discussion of the use of nodes, structures, and bridges, see the Appendix.

Once the model component configuration has been defined using structures and nodes, EcoReliAnT compares input and output node values of structures to identify whether they are in parallel or in series. Equations 14, 15, and 18 are then utilized to solve for the reliability of the model configuration.

A reliability function for the system must also be determined using the system performance metric. This allows for comparison between the model reliability function and the performance reliability function to determine the configuration of components that best model the system performance. Since the method is almost identical, determining

reliability functions for system performance utilizes much of the same code as was developed for microbial populations. In this case, the grouping functionality of the microbial populations was not necessary, as only one performance metric was used to determine the system reliability function. Although the 'saved fit bank' for the system performance is identical to that of the microbial populations, only one reliability function is selected from this 'bank' and used as the function for the system performance.

Comparisons between the model reliability function and the system performance reliability function were made using the sum of squared errors between the two functions calculated at time intervals of 1 day up to a default time of 1000 days. When *k*-out-of-*n* structures are used, each *k* value will result in a different model reliability. Therefore, it was of interest to automate finding the *k* value for each structure that would result in the best fit to the performance reliability function. EcoReliAnT analyzes every combination of possible *k* values in the model system to determine the combination that minimizes the sum of squared errors between the model reliability function and the system reliability function. Finally, EcoReliAnt calculates the reliability importance of each structure in the system, providing information on which structure may be limiting system performance.

## 3.3 Reliability Analysis of External Dataset

To test the ecosystem reliability method and EcoReliAnT software, an external dataset was analyzed. A search of the literature was conducted to identify an engineered ecosystem dataset that provided an adequately long time series of both the system performance and the corresponding functional gene information. We selected a dataset from a study on the effect of trichloroethene (TCE) on microbial community structure and

function in two phenol-fed sequencing batch reactors operated for more than 2 years (Ayala-del-Rio et al. 2004). The 811 days of reactor operation in this study provided an adequately long time series, while the phenol degradation rate from batch assays and RFLP analysis of the phenol hydroxylase gene provided the combination of system performance and functional gene data necessary for reliability analysis. There was instability in the microbial community structure that coincided with changes in reactor performance (Ayala-del-Rio et al. 2004), making the dataset particularly interesting for analysis using our reliability methods.

### 3.3.1 Methodology

System performance information was obtained from Professor Hector Ayala-del-Rio in the form of phenol degradation rates from batch assays (Ayala-del-Rio et al. 2004). Since the intervals between microbial samples were much greater than those between batch assays, and the dates of these measurements do not coincide, the degradation rates were interpolated from the batch assay data to match the microbial sampling dates.

Along with the system performance data, microbial community information was obtained in the form of gel electrophoresis images from RFLP analysis of PCR-amplified phenol hydroxylase genes. These gel electrophoresis images were analyzed using Quantity One® software, Version 4.6.7 (Bio-Rad Laboratories, Hercules, CA), yielding the microbial population time series data for input into EcoReliAnT. Because there was a distinct shift in microbial community structure that occurred during a change in system performance in the phenol-fed sequencing batch reactor from the Ayala-del-Rio study, this data was used for the reliability analysis. One distinction between the gel electrophoresis data and our biofilter data collected using capillary electrophoresis is that the electropherogram peak

areas measured for the gel images were not normalized across samples, and therefore the values represent measures of total abundance. For capillary electrophoresis, there was variation in total abundance due to loading issues, so for these samples I normalized the peak areas. This resulted in information on relative abundance rather than total abundance for the biofilter data.

#### 3.3.1.1 System Reliability Function Determination

A system reliability function was fit to the Weibull model for the interpolated phenol degradation data. The times to failure for the system performance were determined using linear interpolation and a threshold value of 0.195, which is equal to the average degradation rate minus two-fifths of its standard deviation. This value was chosen as it defines spans of good performance near the beginning and end of the time series while excluding the small peak in the middle. Because the small sample size used in this dataset made the chi-square goodness-of-fit test unreliable for determination of the best fitting model, the Weibull model was used for all reliability functions. Right-censoring of times to failure was not considered in the determination of the reliability function, since one of the two times to failure was censored and would have made parameter estimation impossible.

#### 3.3.1.2 Grouping of Microbial Populations

Microbial population dynamics were visually compared to reactor performance to identify populations that were more abundant during periods of good performance and those that were more abundant during periods of poor performance, forming two distinct groups. For the remainder of the analysis, the group of microbial populations that are associated with good performance will be called Group G while those associated with poor

performance will be called Group P. Populations that did not show any correlation between population dynamics and reactor performance were omitted from the analysis.

#### 3.3.1.3 Microbial Population Reliability Function Determination

Times to failure for the group G and P populations were determined using a unique threshold value for each population, based on its average and standard deviation of abundance values. Using this method to define the threshold values ensured that every population would exhibit failures. The times to failure for populations in Group G were determined from spans of presence above the threshold, while times to failure the populations in Group P were determined from spans of absence above the threshold. The times to failure for each microbial population were determined individually from the population dynamics and threshold value, and this failure data was then combined into one dataset for the group from which a group reliability function was determined.

Reliability functions were fit to the failure data using the Weibull model. As with the system reliability model, this dataset provides a small sample size and therefore the Weibull model was used because of its versatility. Right-censoring of times to failure was not accounted for in the determination of the reliability functions, as the high percentage of censored samples would make parameter estimation impossible. To avoid this problem, all right-censored data was considered to be complete data in this thesis. This also provides consistency in the methods used to determine reliability functions between the system performance data and the microbial community data.

#### 3.3.1.4 System Component Configurations

Five different system component configurations were analyzed in this thesis (Figure 14). Cases I and II examined whether the system reliability could be modeled using only the microbial population groups associated with good or poor performance, respectively. This was done by finding a *k*-out-of-*n* structure using only components with the reliability of either Group G or P to describe the performance reliability of the actual system. Cases III, IV, and V examined different *k*-out-of-*n* structure configurations. The number of components in these structures and their reliability functions correspond to the number of members in each of Group G and P and their respective group reliability functions. Case III examined the idea that the system needs both the presence of good populations and the absence of bad populations. This is modeled as a *k*-out-of-*n* structure containing Group G components arranged in series with another *k*-out-of-*n* structure containing Group P components. Case IV examined the idea that the system needs either the presence of good populations or the absence of bad populations. This is modeled as a *k*-out-of-*n* structure containing Group G components arranged in parallel with another *k*-out-of-*n* structure containing Group P components. Finally, Case V examined the idea that the presence of good populations and the absence of bad populations are essentially substitutable. This is modeled as a *k*-out-of-*n* structure containing both Group G and Group P components.

These five component configurations examined different ways in which the microbial populations in the system relate to system performance reliability. Determining which of these configurations best models the reliability of the engineered ecosystem provides a better understanding of how the reliability of engineered ecosystems can be modeled using its functional microbial populations.

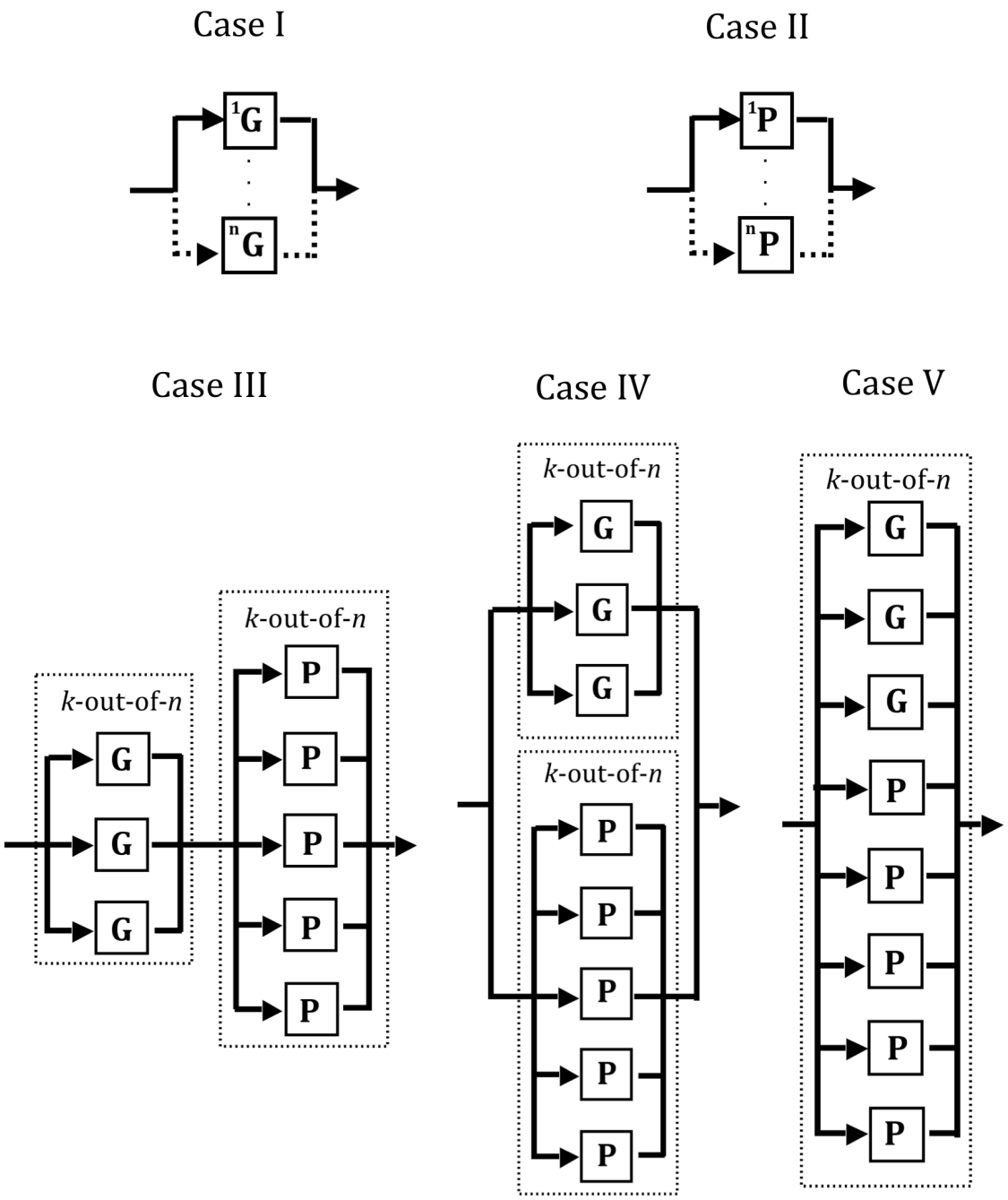



Figure 14. Component configurations analyzed for external dataset.

### 3.3.2 Results

Following the reliability analysis flowchart in Figure 8, reliability functions for the system performance and the microbial populations were determined independently (Steps 1-7). Configurations of microbial components were then compared to the system performance function to determine the best model (Steps 8-10). In this section, I will present the reliability functions for the system performance followed by the reliability functions for the microbial populations. I will then present the results from the comparison of the component configurations (Figure 14) to the system reliability function to determine the best model.

#### 3.3.2.1 System Reliability Function

To calculate the system reliability function, system performance data was obtained in the form of phenol degradation rates. This performance data was then interpolated to coincide with the days that microbial samples were taken. Both the measured and interpolated degradation rates are depicted in Figure 15, and the performance trends are similar between the two. Using a threshold value of 0.195, which is equal to the average phenol degradation rate minus two-fifths of its standard deviation, two 'times to failure' were calculated from the system performance data (Table 2).

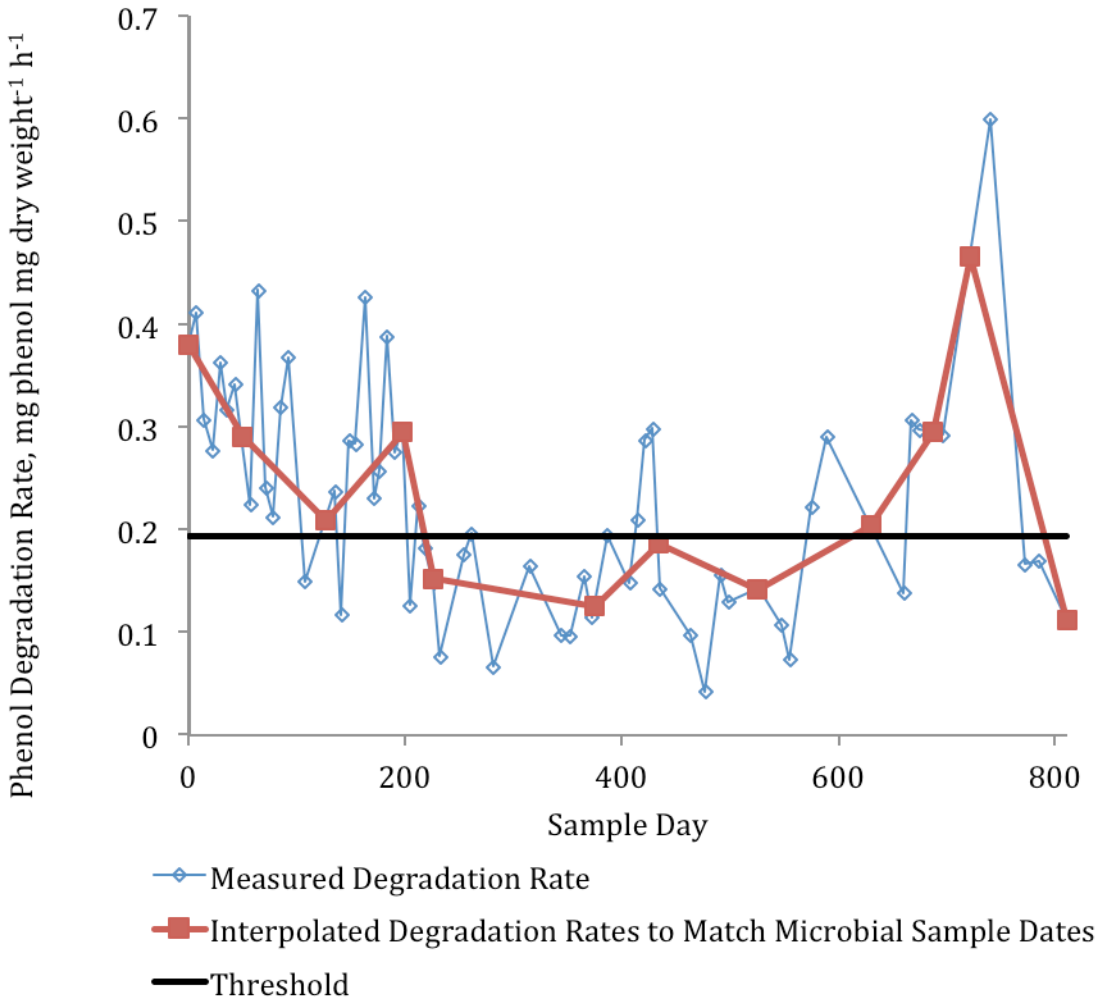



Figure 15. Phenol degradation rate (mg phenol per mg dry weight per hour) was used as the performance metric for the phenol-fed sequencing batch reactor. Data to be used in EcoReliAnT analysis was interpolated from the measured values to correspond with microbial sampling days. 'Times to Failure' for the phenol-fed sequencing batch reactor were determined using a threshold value of 0.1946, the average degradation rate minus two-fifths of its standard deviation.

Table 2. System performance 'time to failure' data based on phenol degradation rate.

| Metric | Threshold Value | Times to Failure (days) |
|---|---|---|
| Phenol Degradation Rate | 0.195 | 217.7<br>176.8 |

From this failure data, maximum likelihood estimates of Weibull reliability model parameters along with their 95% confidence intervals were determined (Table 3), resulting in a reliability function for the system performance (Equation 25). Examination of the reliability and hazard function charts for the system performance (Figure 16) reveal that no system failures occur up to approximately 150 days, at which time the system reliability begins to rapidly decrease. This corresponds to the increase in the hazard rate, or instantaneous failure rate, that occurs at the same time.

Table 3. Maximum likelihood estimates of Weibull reliability model parameters for phenol-fed sequencing batch reactor system performance.

| | Estimated Value | 95% Confidence Interval | |
|---|---|---|---|
| Parameter 1 | 206.56 | 181.95 | 234.50 |
| Parameter 2 | 11.54 | 3.63 | 36.62 |

$$R(t) = e^{-(x/206.56)^{11.54}} \quad (25)$$

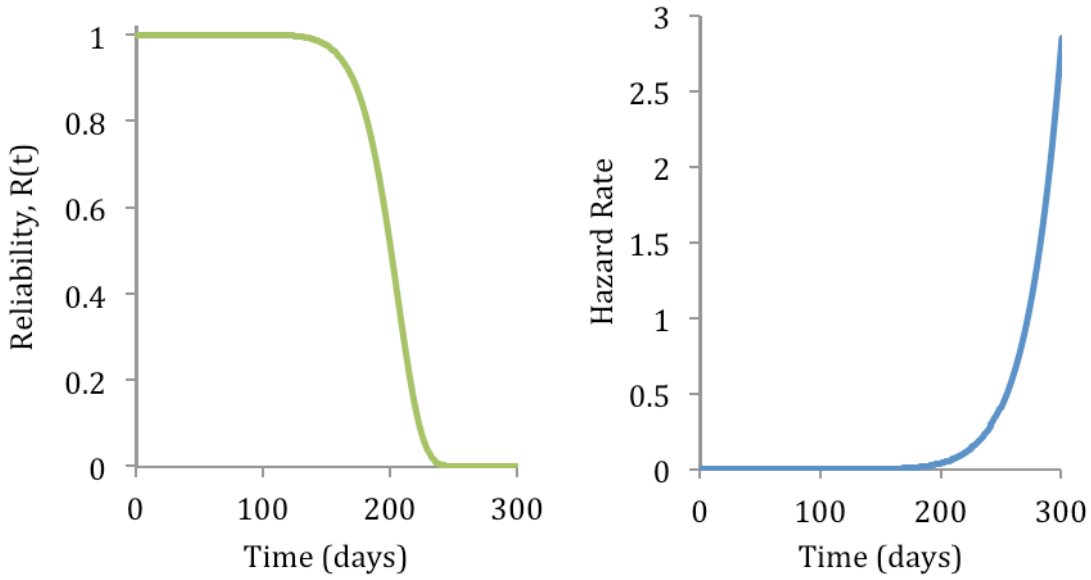


Figure 16. Phenol-fed sequencing batch reactor system performance Weibull reliability function (left) and hazard function (right) using maximum likelihood estimates reveal reliable system performance until approximately 150 days, at which time system failures start to occur.

#### 3.3.2.2 Microbial Population Reliability Functions

Microbial population reliability functions were determined using microbial population dynamics data from RFLP analysis of PCR-amplified phenol hydroxylase genes. Following the reliability analysis flowchart in Figure 8, the next step was to create microbial population groups. Grouping microbial populations by visual comparison between population dynamics and system performance resulted in a group of 3 populations associated with good performance (Figure 17) and a group of 5 populations associated with poor performance (Figure 18). Populations that did not show any correlation between population dynamics and reactor performance were omitted from the analysis. Using linear interpolation and threshold values related to the average and standard deviation of population abundance values (average minus one-half the standard

deviation for Group G and the average minus two-fifths the standard deviation for Group P), two 'times to failure' resulted for each microbial population in Groups G and P (Table 4).

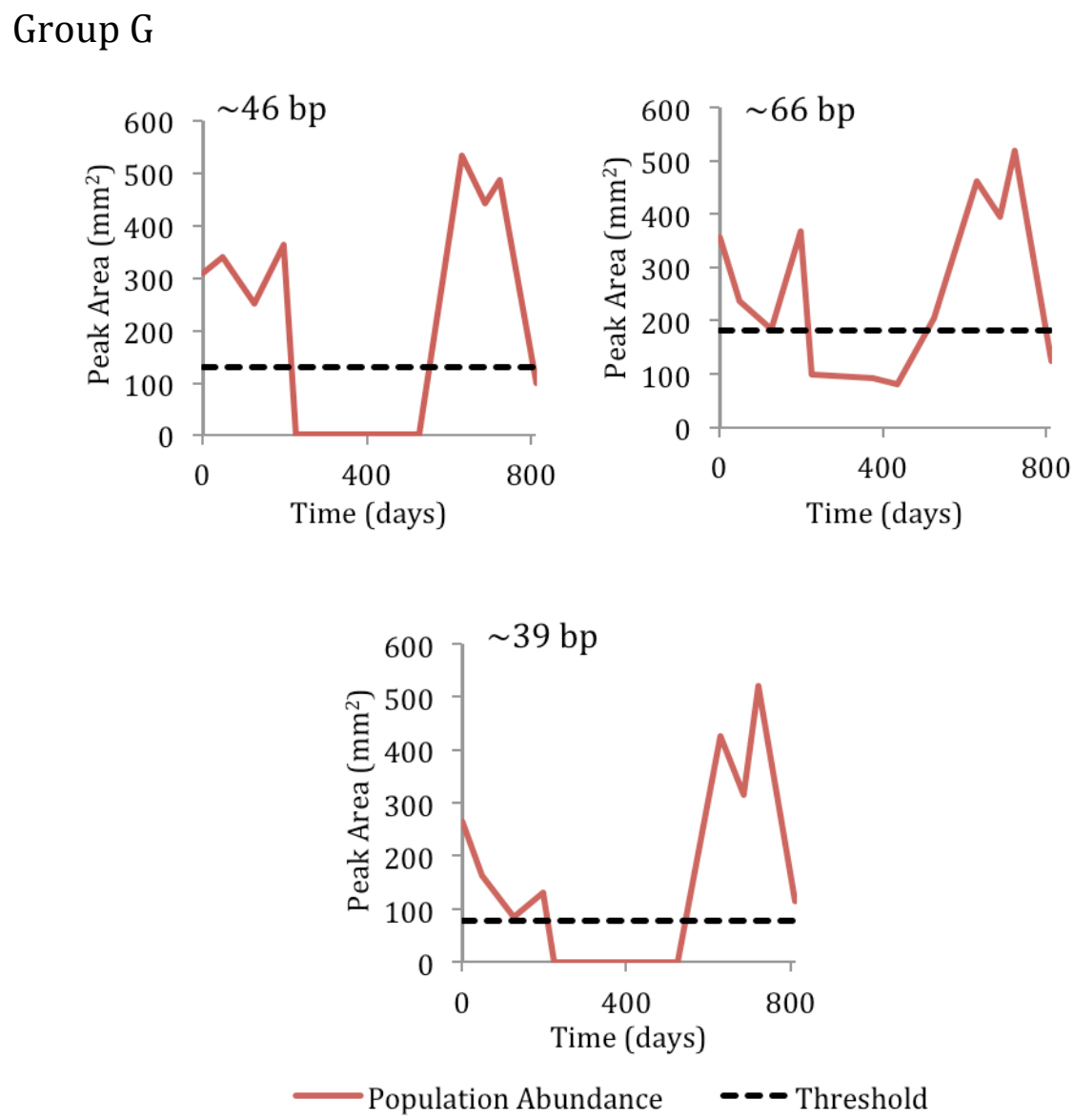



Figure 17. Group G: Microbial populations associated with good performance. 'Times to Failure' for each population were determined from the spans above the threshold value equal to the average minus one-half the standard deviation of the population's abundance values. Population names correspond to the approximate length of nucleic acid fragment in base pairs.

Group P

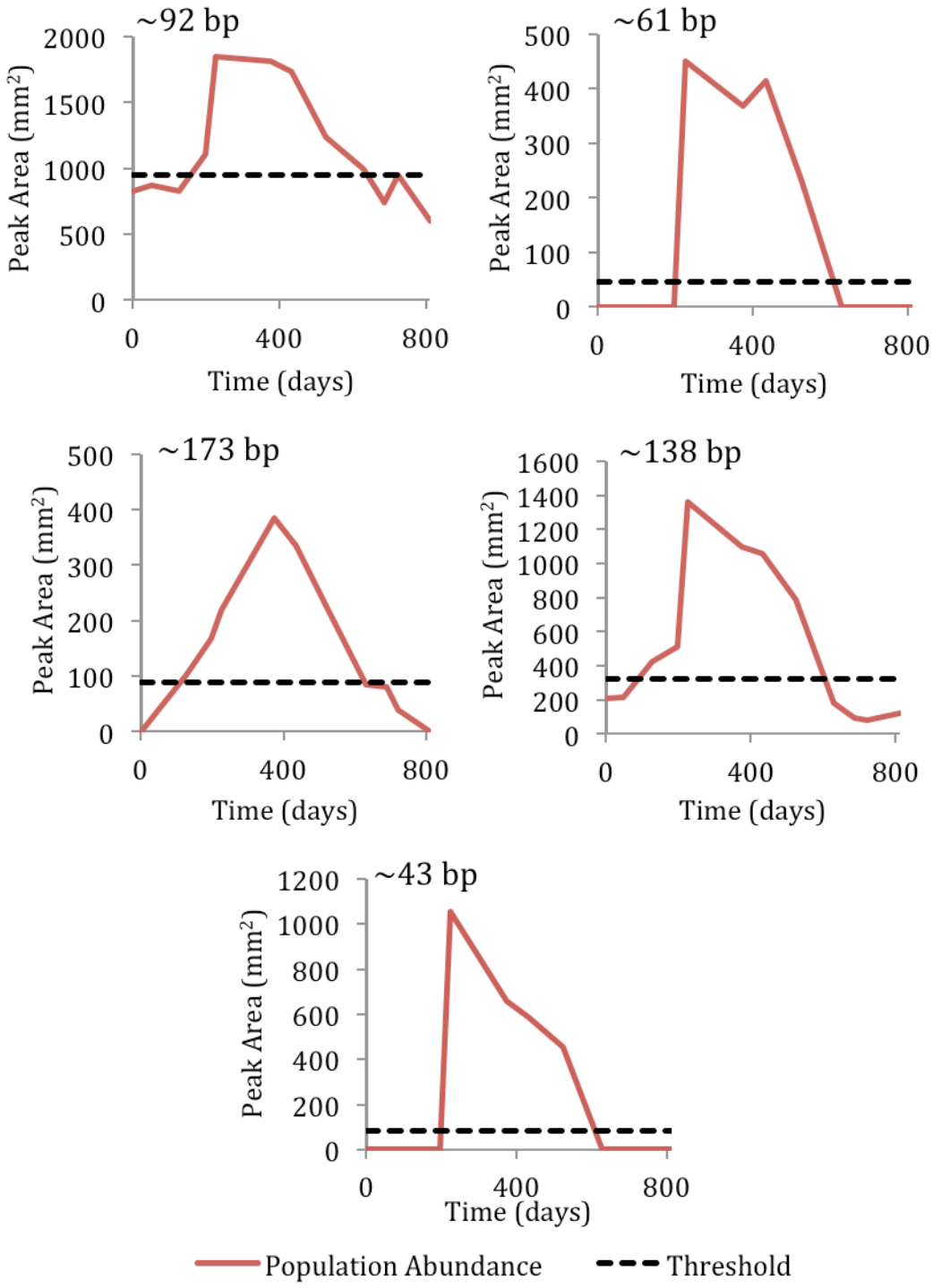



Figure 18. Group P: Microbial populations associated with good performance. 'Times to Failure' for each population were determined from the spans below the threshold value equal to the average minus two-fifths the standard deviation of the population's abundance values. Population names correspond to the approximate length of nucleic acid fragment in base pairs.

Table 4. Microbial population 'time to failure' data. 'Times to failure' were determined for each microbial population based on the specified threshold value. Population names correspond to the approximate length of the nucleic acid fragment.

| Group | Population | Threshold Value | Time to Failure (days) |
|---|---|---|---|
| G | ~66 bp | 182.9 | 209.0 |
| | | | 266.4 |
| | ~46 bp | 131.6 | 215.8 |
| | | | 252.5 |
| | ~39 bp | 79.6 | 217.2 |
| | | | 288.1 |
| P | ~173 bp | 89.0 | 111.5 |
| | | | 184.0 |
| | ~138 bp | 328.1 | 92.3 |
| | | | 207.2 |
| | ~92 bp | 955.0 | 158.8 |
| | | | 172.7 |
| | ~61bp | 47.2 | 200.9 |
| | | | 202.4 |
| | ~43 bp | 83.3 | 200.2 |
| | | | 200.4 |

The failure data was combined for each group and maximum likelihood estimates of Weilbull reliability model parameters along with their 95% confidence intervals were determined separately for Group G and Group P (Table 5), resulting in two reliability functions (Equations 26 and 27). Comparing these reliability functions for each group to the reliability function of the system suggests that the reliability of one of these groups alone does not accurately describe the system reliability, with a sum of squared errors value of 196.75 between the microbial population Group G reliability function and the

system performance reliability function and 59.46 between the system function and the Group P function (Figure 19).

Table 5. Maximum likelihood estimates of Weibull reliability model parameters for microbial population groups were determined from 'time to failure' data.

| | Estimated Value | 95% Confidence Interval | |
|---|---|---|---|
| Group G | | | |
| Parameter 1 | 254.90 | 231.65 | 280.49 |
| Parameter 2 | 8.87 | 4.77 | 16.49 |
| Group P | | | |
| Parameter 1 | 187.36 | 169.22 | 207.46 |
| Parameter 2 | 6.30 | 3.60 | 11.03 |

Group G: $$R(t) = e^{-(x/254.90)^{8.87}} \quad (26)$$

Group P: $$R(t) = e^{-(x/187.36)^{6.30}} \quad (27)$$

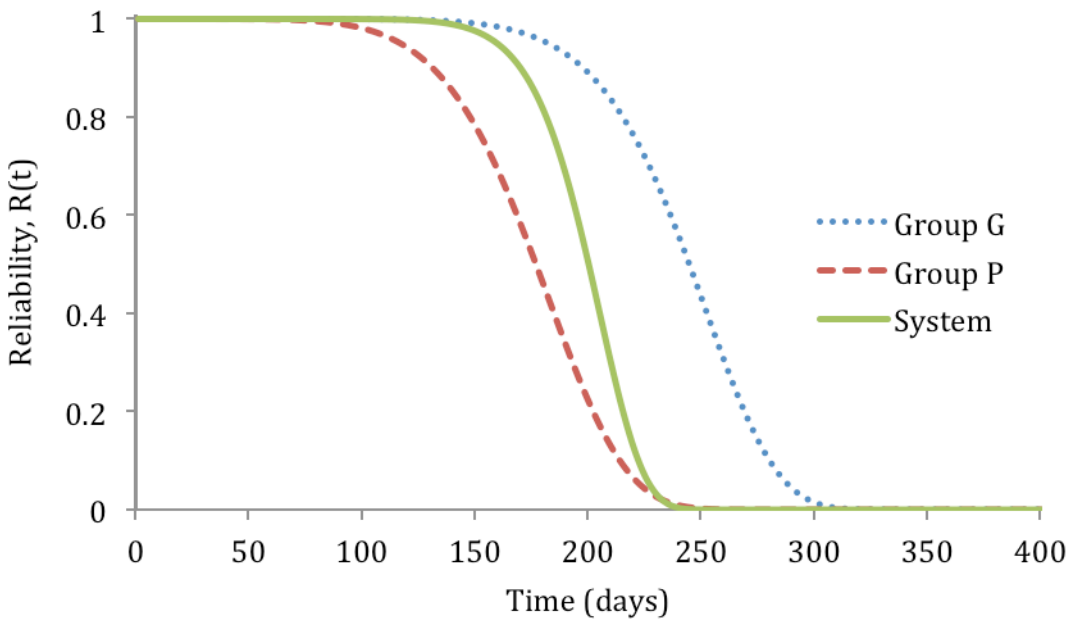



Figure 19. Microbial population group reliability functions and system reliability function. Sum of squared errors value between the system performance reliability function and reliability functions for Groups G and P are 59.46 and 196.75, respectively.

#### 3.3.2.3 Analysis of Component System Configurations

In addition to the comparison between the group reliability functions and the system performance function (Figure 19), 5 different component configurations (Figure 14) were analyzed to determine the best fitting model.

Case I examined the possibility that the system reliability could be modeled by a *k*-out-of-*n* structure using only components with the reliability of Group G. Analyzing this configuration of components with *n* values ranging from 1 to 25, the *k* and *n* values that most closely modeled the system performance were 14 and 15, respectively. This configuration results in a calculated sum of squared errors value of 0.71 between the model reliability function and the system reliability function. When the number of components in Group G (3) was used as the *n* value, the best fitting model was a 1-out-of-3 structure. This model resulted in a 34.86 sum of squared errors value between the model reliability function and the system reliability function.

Similarly, Case II examined the possibility that the system reliability could be modeled by a *k*-out-of-*n* structure using only components with the reliability of Group P. This configuration of components was also analyzed with *n* values ranging from 1 to 25. The *k* and *n* values that most closely modeled the system performance were 1 and 3, respectively. This configuration results in a calculated sum of squared errors value of 1.04 between the model reliability function and the system reliability function. When the number of components in Group P (5) was used as the *n* value, the best fitting model was a 2-out-of-5 structure. This model resulted in a 9.98 sum of squared errors value between the model reliability function and the system reliability function.

In Case III, consisting of a Group G *k*-out-of-3 structure in series with a Group P *k*-out-of-5 structure, the best fit *k* values were determined to be 3 and 1, respectively. This means that the reliability of this model best matches the performance reliability of the actual system when all three of the populations from Group G are present while one of the populations from Group P is absent. This system configuration then defines failure as either one of the populations from Group G no longer being present or all five of the populations from Group P being present. This configuration resulted in a calculated sum of squared errors value of 0.07 between the model reliability function and the system reliability function.

In Case IV, consisting of a Group G *k*-out-of-3 structure in parallel with a Group P *k*-out-of-5 structure, the best fit *k* values were determined to be 3 and 5, respectively. Therefore, the reliability of this system best matches the reliability of the real system when all three of the populations from Group G are present or all five of the populations from Group P are absent. This system configuration then defines failure as one of the populations from Group G no longer being present and one of the populations from Group P being present. This configuration resulted in a calculated sum of squared errors value of 34.96 between the model reliability function and the system reliability function.

Analysis of Case V, consisting of a *k*-out-of-*n* structure containing 3 Group G components and 5 Group P components, resulted in a best fit *k* value of 4. Therefore, the reliability of this system best matches the reliability of the real system when the number of populations from Group G that are present and the number of populations from Group P that are absent add up to 4. This configuration resulted in a calculated sum of squared

errors value of 4.42 between the model reliability function and the system reliability function

Of the cases analyzed, Case III consisting of a 3-out-of-3 Group G structure in series with a 1-out-of-5 Group P structure most closely matched the reliability of the actual system performance (Table 6, Figure 20). This system configuration, which defines failure as either one of the populations from Group G no longer being present or all five of the populations from Group P being present, resulting in a reliability function almost identical to the function determine for the actual system. Additionally, all of the cases analyzed resulted in reliability functions more closely matching the system performance than the reliability functions of Group G and P alone.

Table 6. Sum of squared errors values between model and system performance reliability functions. For Cases I and II, the best fitting *k*-out-of-*n* structure with and without *n* being fixed at the number of group members is shown.

| | Case | SSE |
|---|---|---|
| Individual Reliability Functions | Group G | 196.75 |
| | Group P | 59.46 |
| Component Configurations | Case I (any $n$ <25) | 0.71 |
| | Case I ($n$ = 3) | 34.86 |
| | Case II (any $n$ <25) | 1.04 |
| | Case II ($n$ = 5) | 9.98 |
| | Case III | 0.07 |
| | Case IV | 34.96 |
| | Case V | 4. |

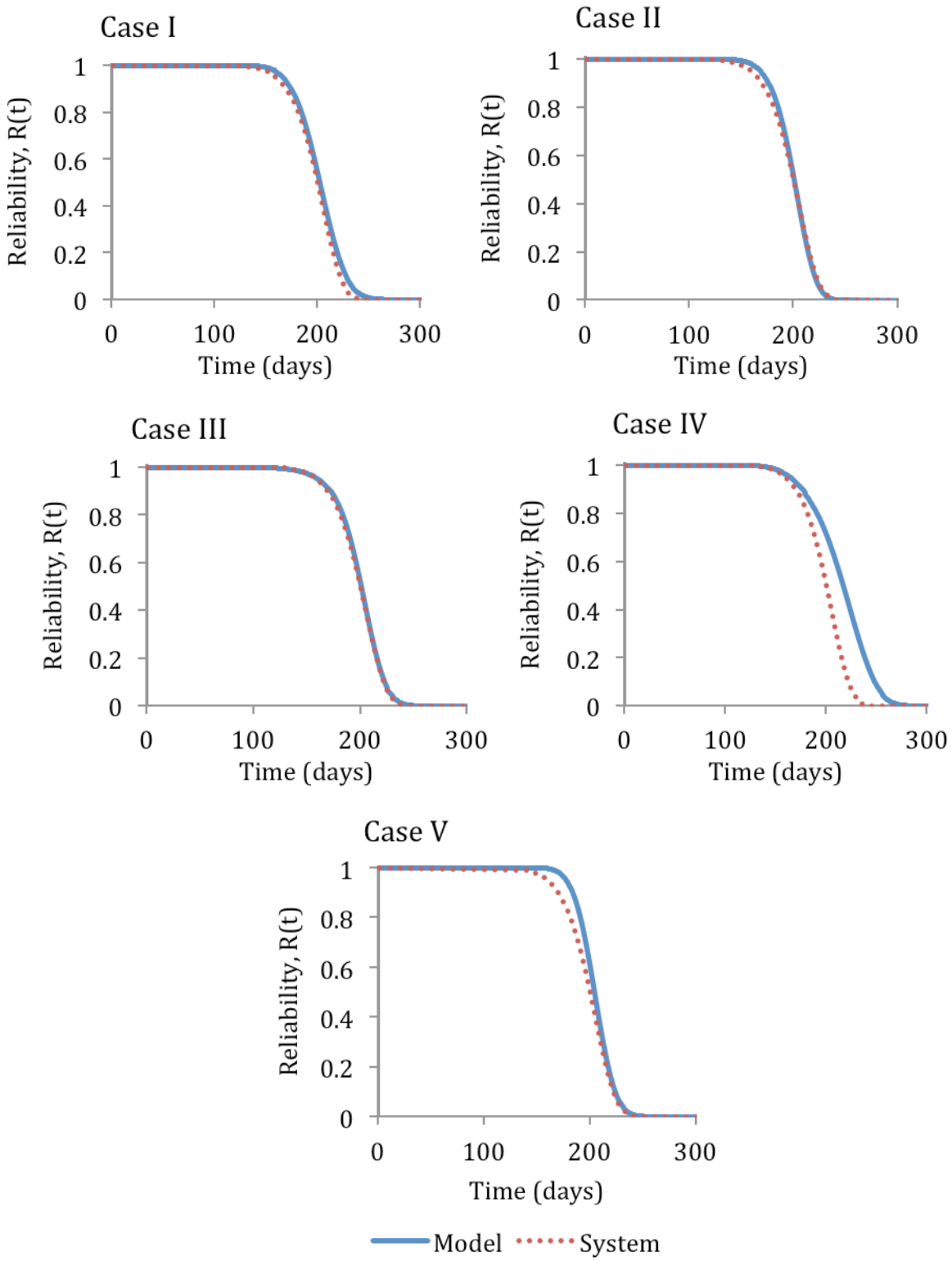



Figure 20. Comparison of model reliability functions and phenol-fed reactor system performance reliability function for the five component configurations analyzed.

## 3.4 Field Biofilter Analysis

After testing the ecosystem reliability analysis method and EcoReliAnT software on a relatively simple dataset from the literature, the next step was to analyze data from a field biofilter. Of the functional genes responsible for the stepwise reduction of nitrate to dinitrogen gas (Figure 1), only *nosZ* has been quantified in our samples thus far. Therefore, in this thesis the analysis of the field biofilters is in essence reduced to a one-step system consisting of the *nosZ* populations. ARISA and Fungal ARISA data exist for the field biofilters, but have not yet been included because they are not directly linked to denitrification activity. How these populations should be included in the system model has not yet been determined and requires further examination.

### 3.4.1 Methodology

For the reliability analysis in this thesis, I focused on data from the FP07 biofilter in Decatur, Illinois from 1/5/09 to 5/19/09. This selection was based on the availability of performance and microbial data during a period of continuous flow and the existence of variation in system performance.

Nitrate removal efficiency was used as the system performance metric for the denitrifying biofilter reliability analysis. This data was obtained from influent and effluent grab samples collected by Dr. Richard Cooke and his laboratory (Dept. of Agricultural and Biological Engineering, University of Illinois at Urbana-Champaign). The nitrate concentrations in these samples were measured by Agricultural and Biological Engineering Water Quality Laboratory (University of Illinois, Urbana, IL) using EPA Method 353.1. Since the intervals between microbial samples and nitrate samples varied and the dates of these

measurements do not coincide, the nitrate removal efficiency values used in EcoReliAnT were interpolated from the measured data to match the microbial sampling dates.

Microbial community information was obtained from *nosZ* T-RFLP analysis of the community in woodchip samples taken from sampling ports in the biofilters (Andrus 2010; Porter 2011). The FP07 biofilter contains five sampling ports: 3 shallow ports and 2 deep ports (Figure 21). Woodchip samples were collected from all 5 ports at approximate intervals of 2 weeks during the time period in this analysis. DNA extraction, purification, and microbial community structure determination were undertaken by Malia Andrus (2010) and Matt Porter (2011). Denitrifying bacteria possessing the nitrous oxide reductase (*nosZ*) gene were analyzed with terminal restriction fragment length polymorphism analysis using the *nosZ* F-1181 forward primer and the *nosZ* R-1880 reverse primer (Rich et al. 2003). DNA fragments from these techniques were analyzed using denaturing capillary electrophoresis (Porter 2011). Fragment length data was then analyzed using using GeneMarker software, version 1.95 (SoftGenetics, State College, PA). Fragments between 100-700 bp were analyzed and samples with poor size calling or overall low fluorescence were excluded. Relative fluorescence of *nosZ* terminal restriction fragments (TRFs) was determined separately for each restriction digest. Relative fluorescence values for input into EcoReliAnT were averaged across all ports in the biofilter for every time point. While there was a lot of variation in each population's relative abundance values among the ports for a given date, with relative standard deviation values ranging from 67% to 224%, I felt that averaging the values provided a dataset that was more representative of the community in entire biofilter.

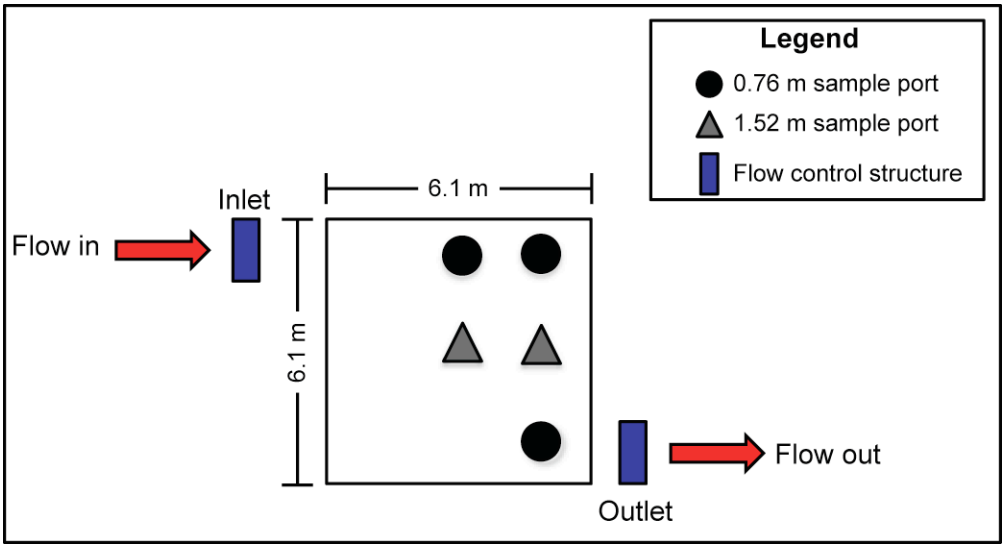


Figure 21. Schematic of the FP07 biofilter with flow direction, sampling port locations, and biofilter dimensions indicated. From Porter, 2011.

#### 3.4.1.1 System Reliability Function Determination

A system reliability function was fit to the Weibull model using percent nitrate removal as the system performance metric. The times to failure for the system performance were determined using a threshold value of 0.73, which is equal to the average degradation rate minus two-fifths of its standard deviation. As with the external dataset analysis, this dataset provided a small sample size and therefore the Weibull model was used because of its versatility. Right-censoring of times to failure was not accounted for in the determination of the reliability function, since two of the three times to failure were censored so parameter estimation would have been impossible.

#### 3.4.1.2 Grouping of Microbial Populations

Microbial population reliability functions were determined from population dynamics data obtained through T-RFLP analysis of the nitrous oxide reductase (*nosZ*) gene. Following the reliability analysis flowchart in Figure 8, the next step in the

determination of microbial reliability was to create groups to increase the number of observed failures. Unlike the external dataset, creating groups of microbial populations associated with good and poor performance by visual comparison was cumbersome, as there were 117 populations present in this dataset. In order to create groups, a modified version of the Local Similarity Analysis method was used (Ruan et al. 2006). This method compares each microbial population curve to the system performance, providing a test statistic (LS Score) and p-value. I modified the method to compound the test statistic for the entire time series. Groups of microbial populations both positively and negatively associated with system performance were identified using LS score and p-value thresholds of 0.2 and 0.1, respectively. As with the external dataset analysis, the group of microbial populations that are associated with good performance will be called Group G while those associated with poor performance will be called Group P. This method provided a substantial reduction in populations, from 117 to 8, making the reliability analysis feasible.

#### 3.4.1.3 Reliability Functions for Microbial Populations

Times to failure for the populations were determined using a unique threshold value for each population based on its average and standard deviation of relative abundance values. The same method was used to determine the threshold values of all microbial populations to remain consistent across the entire dataset. The times to failure for populations in Group G were determined from spans of absence below the threshold while times to failure the populations in Group P were determined from spans of presence above the threshold. Using this method to define the threshold values ensured that every population would exhibit failures. The times to failure for each microbial population were determined individually from its population dynamics and threshold value, and this failure

data was then combined into one dataset for the group from which a group reliability function was determined.

Reliability functions were fit to the failure data using the Weibull model. As with the system reliability model, this dataset provides a small sample size and therefore the Weibull model is used because of its versatility. Right-censoring of times to failure was not accounted for in the determination of the reliability functions, providing consistency in the methods used to determine reliability functions between the system performance data and the microbial community data.

#### 3.4.1.4 System Component Configurations

Following the external dataset analysis procedure discussed earlier, five different system component configurations were analyzed for the field denitrifying biofilter data (Figure 22). These configurations were identical to those from the external dataset analysis except for the number of components in the *k*-out-of-*n* structures (Cases III, IV, and V), which corresponded to the number of members in Groups G and P.

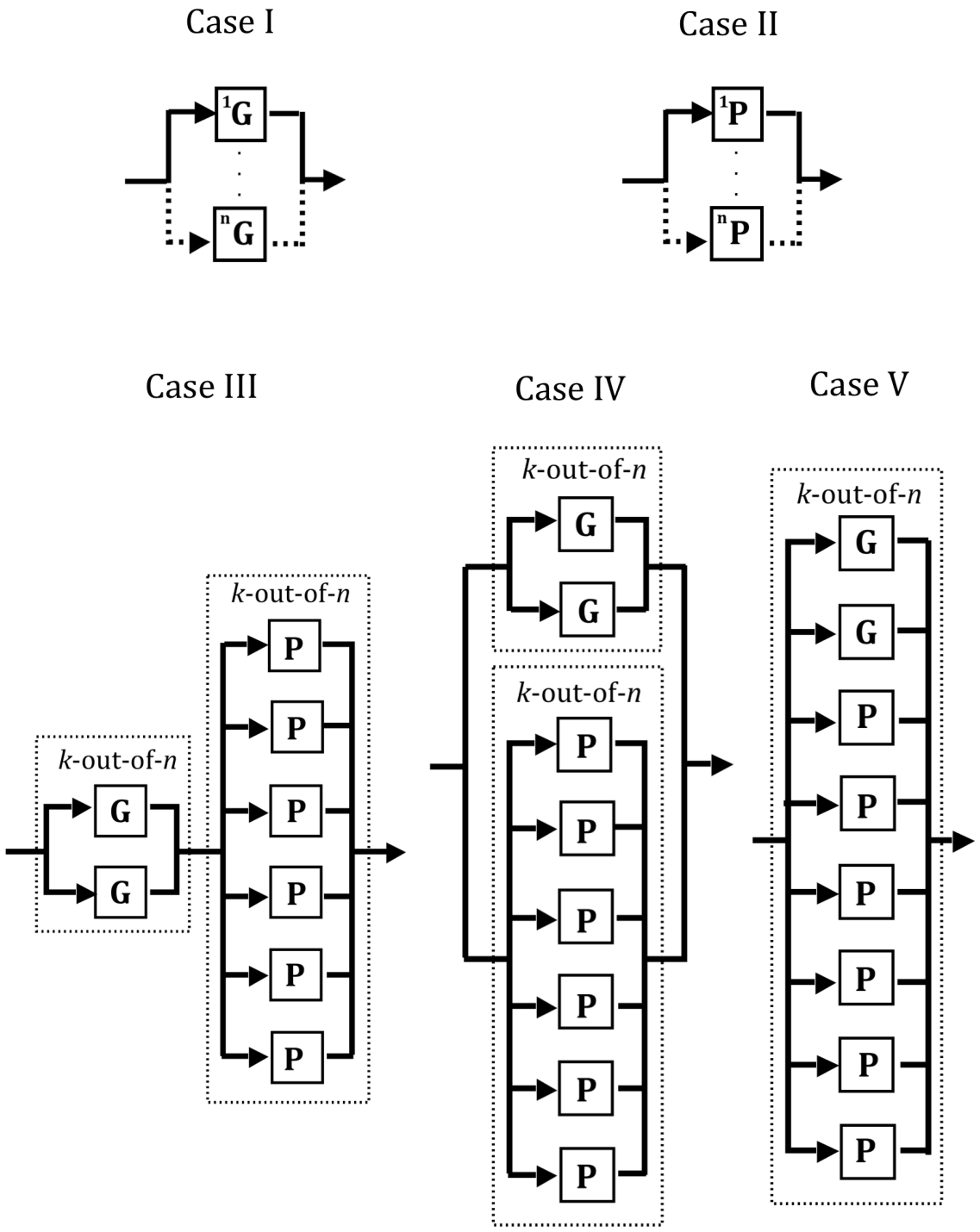



Figure 22. Reliability analysis component configurations used in analysis of FP07 field denitrifying biofilter data.

### 3.4.2 Results

Following the reliability analysis flowchart in Figure 8, reliability functions for the system performance and the microbial populations are determined independently (Steps 1-7) Configurations of microbial components are then compared to the system performance function to determine the best model (Steps 8-10). As with the external dataset results, I will first present the reliability functions for the system performance and microbial populations. I will then present the results from the comparison of the component configurations (Figure 22) to the system reliability function to determine the best model.

#### 3.4.2.1 System Reliability Function

The system reliability function for the FP07 biofilter was determined from percent nitrate removal data. This performance data was linearly interpolated to coincide with the days that microbial samples were taken. Both the measured and interpolated degradation rates are depicted in Figure 23, and the performance trends are similar between the two. Using a threshold value of 0.73, which is equal to the average nitrate removal rate minus two-fifths of its standard deviation, three 'times to failure' were observed in the system performance data (Table 7).

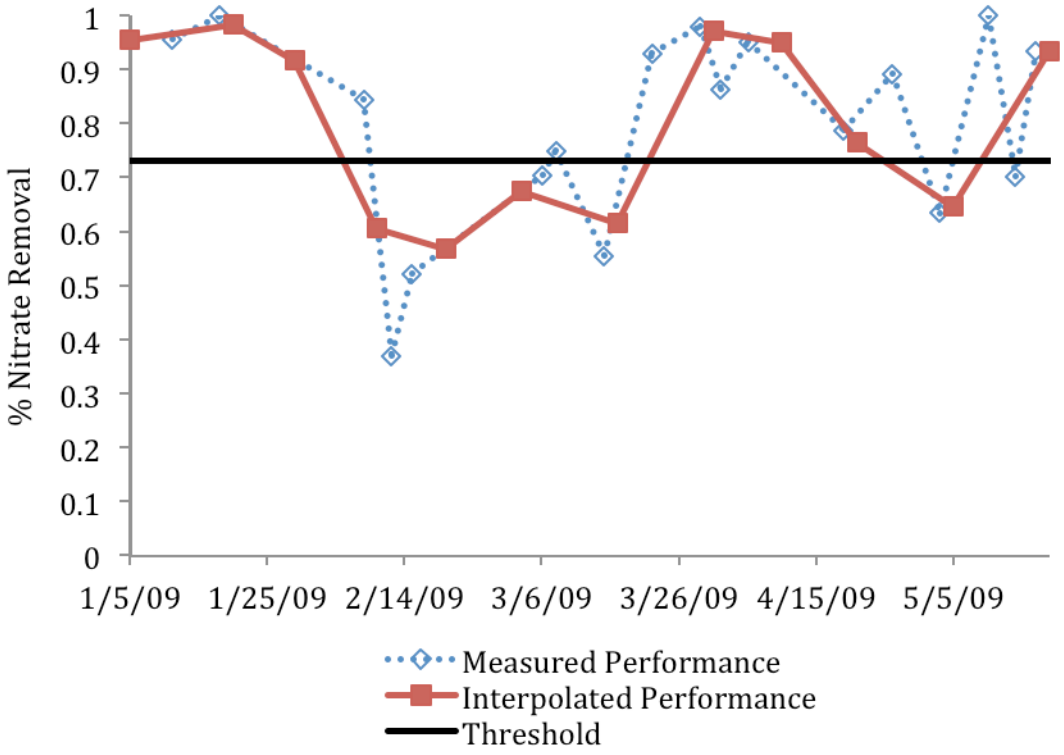



Figure 23. Percent nitrate removal is used as the performance metric for the FP07 field denitrifying biofilter. Data to be used in EcoReliAnT analysis was interpolated from the measured values provided to correspond with microbial sampling days. 'Times to Failure' were determined using a threshold value of 0.73, the average degradation rate minus two-fifths of its standard deviation.

Table 7. System performance 'Time to Failure' data.

| Metric | Threshold Value | Times to Failure (days) |
|---|---|---|
| Percent Nitrate Removal | 0.73 | 31.2<br>34.3<br>9.8 |

Maximum likelihood estimates of Weibull reliability model parameters along with their 95% confidence intervals were then determined from this failure data (Table 8), resulting in a reliability function for the system performance (Equation 28). Examination of the reliability and hazard function charts for the system performance (Figure 24) reveal that system failures begin to occur fairly quickly, with the system reliability beginning to

rapidly decrease around 15 days. This corresponds to the increase in the hazard rate, or instantaneous failure rate, that occurs at the same time.

Table 8. Maximum Likelihood Estimates of Weibull Reliability Model Parameters for System Performance.

| | Estimated Value | 95% Confidence Interval | |
|---|---|---|---|
| Parameter 1 | 28.29 | 18.02 | 44.43 |
| Parameter 2 | 2.62 | 0.94 | 7.29 |

$$R(t) = e^{-(x/28.29)^{2.62}} \quad (28)$$

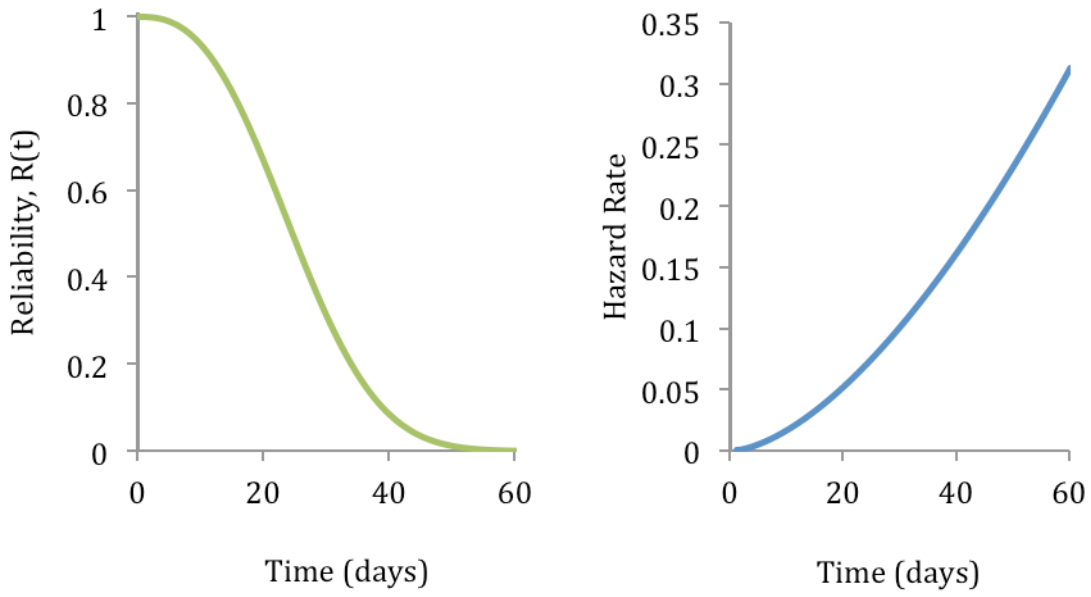



Figure 24. FP07 field denitrifying biofilter Weibull reliability function (left) and hazard function (right) using maximum likelihood estimates.

### 3.4.2.2 Microbial Population Reliability Functions

Microbial population reliability functions were determined from population dynamics obtained through T-RFLP analysis of the *nosZ* gene. Local Similarity Analysis of

the population dynamics and system performance data resulted in groups of populations positively and negatively correlated with system performance (Table 9). 'Times to failure' for each of these populations were determined using linear interpolation and threshold values of the average minus two-fifths of the standard deviation of the population abundance (Figure 25, Figure 26, Table 10).

Table 9. Local Similarity Analysis results comparing microbial population dynamics to biofilter performance data. Groups of microbial populations associated with good and poor performance were created using LS Score and p-value thresholds of 0.2 and 0.1, respectively.

| Population | LS Score | p-value | Positive or Negative Correlation |
|---|---|---|---|
| Alu-435 | 0.483 | 0.010 | Negative |
| Alu-440 | 0.296 | 0.010 | Negative |
| Hha-101 | 0.351 | 0.021 | Positive |
| Alu-318 | 0.282 | 0.023 | Negative |
| Hha-102 | 0.283 | 0.048 | Negative |
| Alu-173 | 0.244 | 0.065 | Negative |
| Alu-226 | 0.280 | 0.093 | Negative |
| Alu-465 | 0.326 | 0.098 | Positive |

Group G

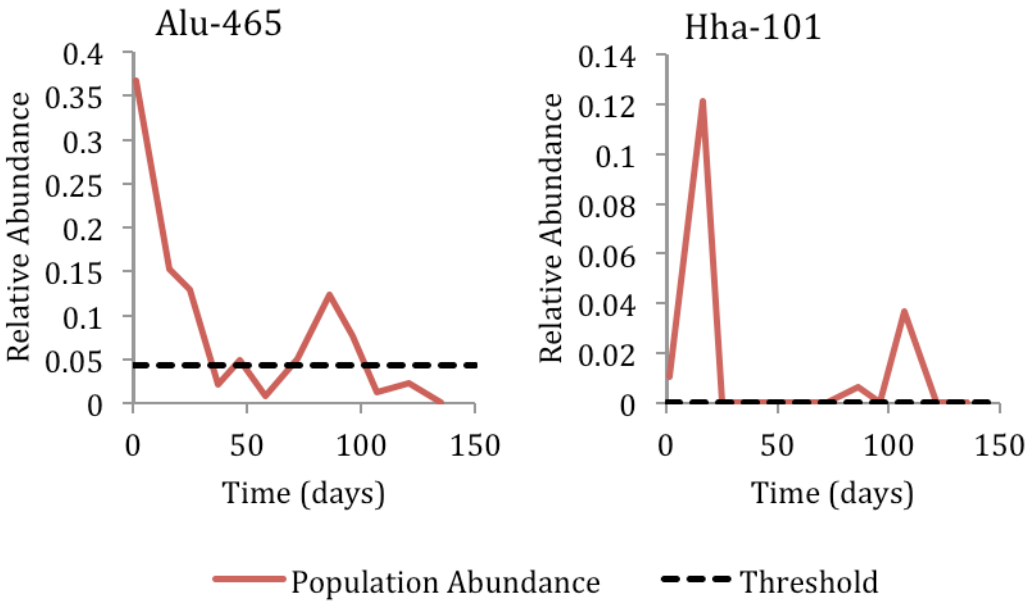



Figure 25. Group G: Microbial populations correlated to good performance. 'Times to Failure' for each population were determined from the spans above the threshold value equal to the average minus two-fifths the standard deviation of the population's abundance values.

Group P

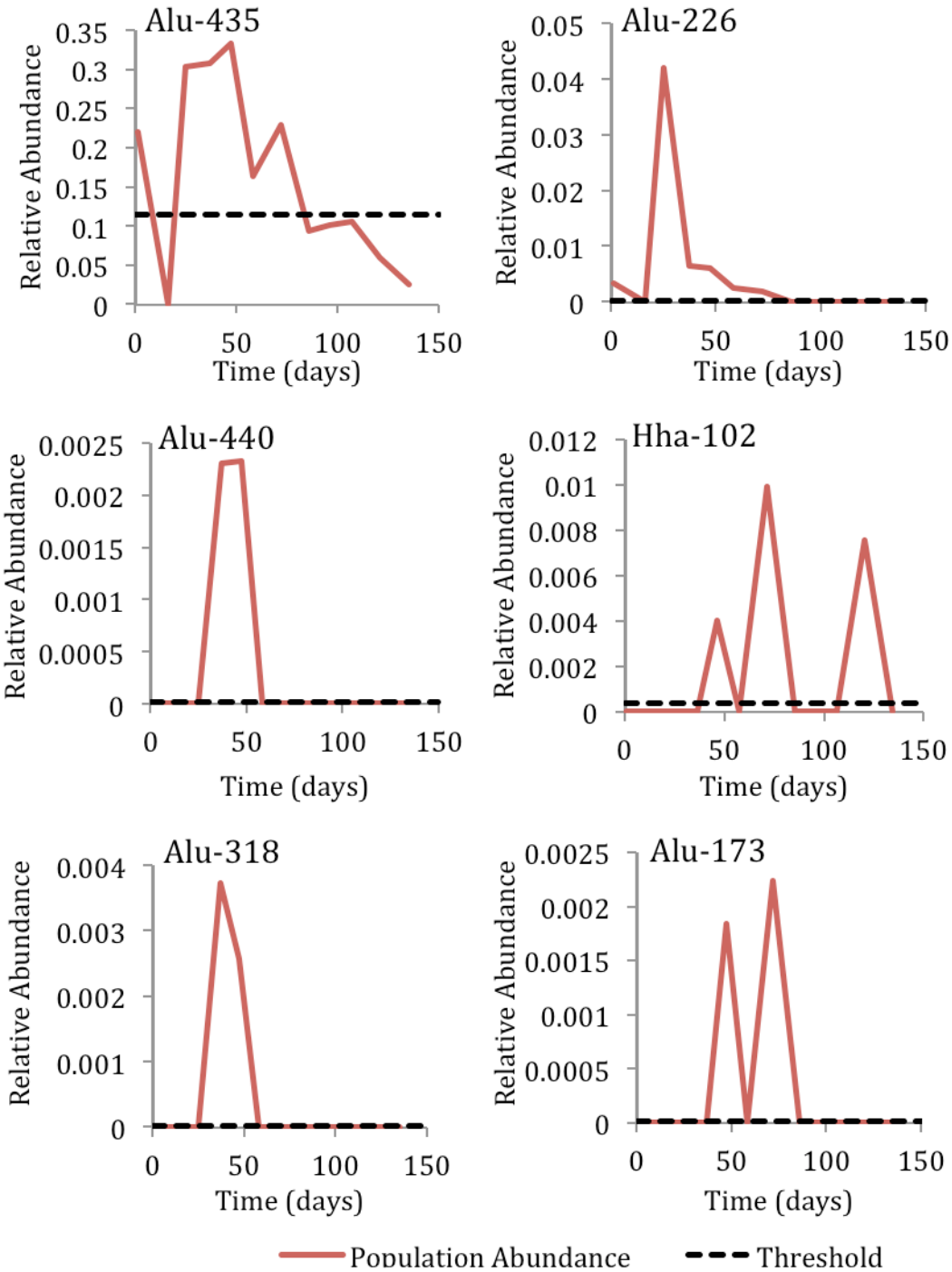



Figure 26. Group P: Microbial populations correlated to poor performance. 'Times to Failure' for each population were determined from the spans below the threshold value equal to the average minus two-fifths the standard deviation of the population's abundance values.

Table 10. Microbial population 'time to failure' data. 'Times to failure' were determined for each microbial population based on the specified threshold value. Population names correspond to the restriction enzyme used and approximate length of the nucleic acid fragment in base pairs.

| Group | Population | Threshold Value | Time to Failure (days) |
|---|---|---|---|
| G | Hha-101 | 3.78e-04 | 24.0 |
| | | | 22.5 |
| | | | 24.7 |
| | Alu-465 | 0.0428 | 33.6 |
| | | | 3.6 |
| | | | 32.0 |
| P | Alu-435 | 0.115 | 11.3 |
| | | | 51.3 |
| | Alu-440 | 2.55e-05 | 24.1 |
| | | | 77.1 |
| | Alu-318 | 2.45e-05 | 24.1 |
| | | | 77.1 |
| | Hha-102 | 3.97e-04 | 37.0 |
| | | | 1.7 |
| | | | 22.3 |
| | | | 0.7 |
| | Alu-173 | 2.06e-05 | 36.1 |
| | | | 0.3 |
| | | | 49.1 |
| | Alu-226 | 4.38e-04 | 2.1 |
| | | | 52.4 |

The failure data was combined for each group and maximum likelihood estimates of Weilbull reliability model parameters along with their 95% confidence intervals were determined from this failure data for Group G and Group P (Table 11), resulting in reliability functions for each group of microbial populations (Equations 29 and 30). Comparing the reliability functions of each group to the reliability function of the system shows that Group G has a very similar reliability function to the system, while the shape of the reliability function for Group P is much different (Figure 27). The sum of squared errors

between the microbial population group reliability functions and the system reliability functions for Groups G and P are 1.28 and 24.13, respectively.

Table 11. Maximum likelihood estimates of Weibull reliability model parameters for microbial population groups were determined from 'time to failure' data.

| | Estimated Value | 95% Confidence Interval | |
|---|---|---|---|
| Group G | | | |
| Parameter 1 | 25.94 | 18.64 | 36.11 |
| Parameter 2 | 2.49 | 1.19 | 5.17 |
| Group P | | | |
| Parameter 1 | 29.31 | 15.85 | 54.19 |
| Parameter 2 | 0.86 | 0.55 | 1.33 |

Group G: $$R(t) = e^{-(x/25.94)^{2.49}} \quad (29)$$

Group P: $$R(t) = e^{-(x/29.31)^{0.86}} \quad (30)$$

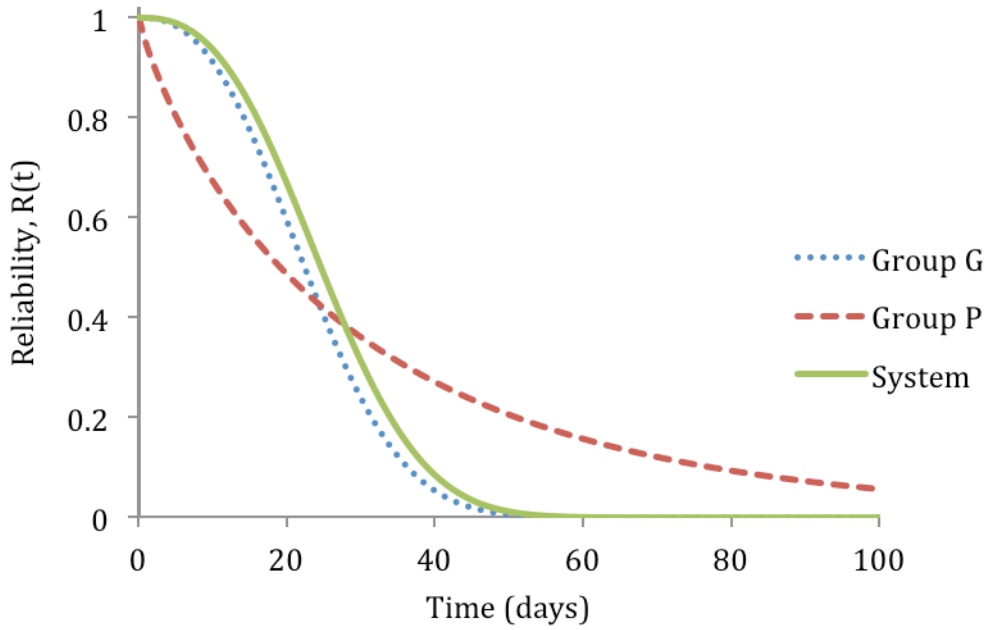



Figure 27. Comparison of individual microbial population group reliability functions and system reliability function shows that the reliability of Group G alone fairly accurately models the system reliability, while Group P does not. Resulting sum of squared errors between the system reliability function and the reliability functions for Group G and P are 1.28 and 24.13, respectively.

#### 3.4.2.3 Analysis of Component System Configurations

In addition to the comparison between the group reliability functions and the system performance function (Figure 27), 5 different component configurations (Figure 22) were analyzed to determine the best fitting model.

Case I examined the possibility that the system reliability could be modeled by a *k*-out-of-*n* structure using only components with the reliability of Group G. Analyzing this configuration of components with *n* values ranging from 1 to 25, the *k* and *n* values that most closely modeled the system performance were 1 and 1, respectively. This configuration results in a calculated sum of squared errors value of 1.28 between the model reliability function and the system reliability function. When the number of components in

Group G (2) was used as the *n* value, the best fitting model was a 1-out-of-2 structure. This model resulted in a 4.26 sum of squared errors value between the model reliability function and the system reliability function.

Similarly, Case II examined the possibility that the system reliability could be modeled by a *k*-out-of-*n* structure using only components with the reliability of Group P. This configuration of components was also analyzed with *n* values ranging from 1 to 25. The *k* and *n* values that most closely modeled the system performance were 6 and 13, respectively. This configuration results in a calculated sum of squared errors value of 0.24 between the model reliability function and the system reliability function. When the number of components in Group P (6) was used as the *n* value, the best fitting model was a 3-out-of-6 structure. This model resulted in a 3.31 sum of squared errors value between the model reliability function and the system reliability function.

In Case III, consisting of a Group G *k*-out-of-2 structure in series with a Group P *k*-out-of-6 structure, the best fit *k* values were determined to be 1 and 2, respectively. Therefore, the reliability of this model best matches the reliability of the system performance when one of the populations from Group G are present and two of the populations from Group P are absent. This system configuration then defines failure as either both of the populations from Group G no longer being present or five of the six populations from Group P being present. This configuration resulted in a calculated sum of squared errors value of 0.70 between the model reliability function and the system reliability function.

In Case IV, consisting of a Group G *k*-out-of-2 structure in parallel with a Group P *k*-out-of-6 structure, the best fit *k* values were determined to be 2 and 4, respectively.

Therefore, the reliability of this model best matches the reliability of the system performance when both of the populations from Group G are present or four of the populations from Group P are absent. This system configuration then defines failure as one of the populations from Group G no longer being present and three of the populations from Group P being present. This configuration resulted in a calculated sum of squared errors value of 3.09 between the model reliability function and the system reliability function.

Analysis of Case V, consisting of a *k*-out-of-*n* structure containing 2 Group G components and 6 Group P components each possessing,  resulted in a best fit *k* value of 4. Therefore, the reliability of this model best matches the reliability of the system performance when the number of populations from Group G that are present and the number of populations from Group P that are absent add up to 4. This configuration resulted in a calculated sum of squared errors value of 0.53 between the model reliability function and the system reliability function.

Of the cases analyzed, Case II, consisting of a 6-out-of-13 system composed solely of members of the microbial populations positively correlated with poor system performance, most closely matched the reliability of the actual system (Table 12, Figure 28). While Group G alone modeled the system performance well, Cases II, III, and V all resulted in reliability functions more closely matching the system performance than the reliability functions of either Group G and P.

Table 12. Sum of squared errors values between model and system performance reliability functions. For Cases I and II, the best fitting *k*-out-of-*n* structure with and without *n* being fixed at the number of group members is shown.

| | Case | SSE |
|---|---|---|
| Individual Reliability Functions | Group G | 1.28 |
| | Group P | 24.13 |
| Component Configurations | Case I (any $n$ <25) | 1.28 |
| | Case I ($n$ = 2) | 4.26 |
| | Case II (any $n$ <25) | 0.24 |
| | Case II ($n$ = 6) | 3.31 |
| | Case III | 0.70 |
| | Case IV | 3.09 |
| | Case V | 0.53 |

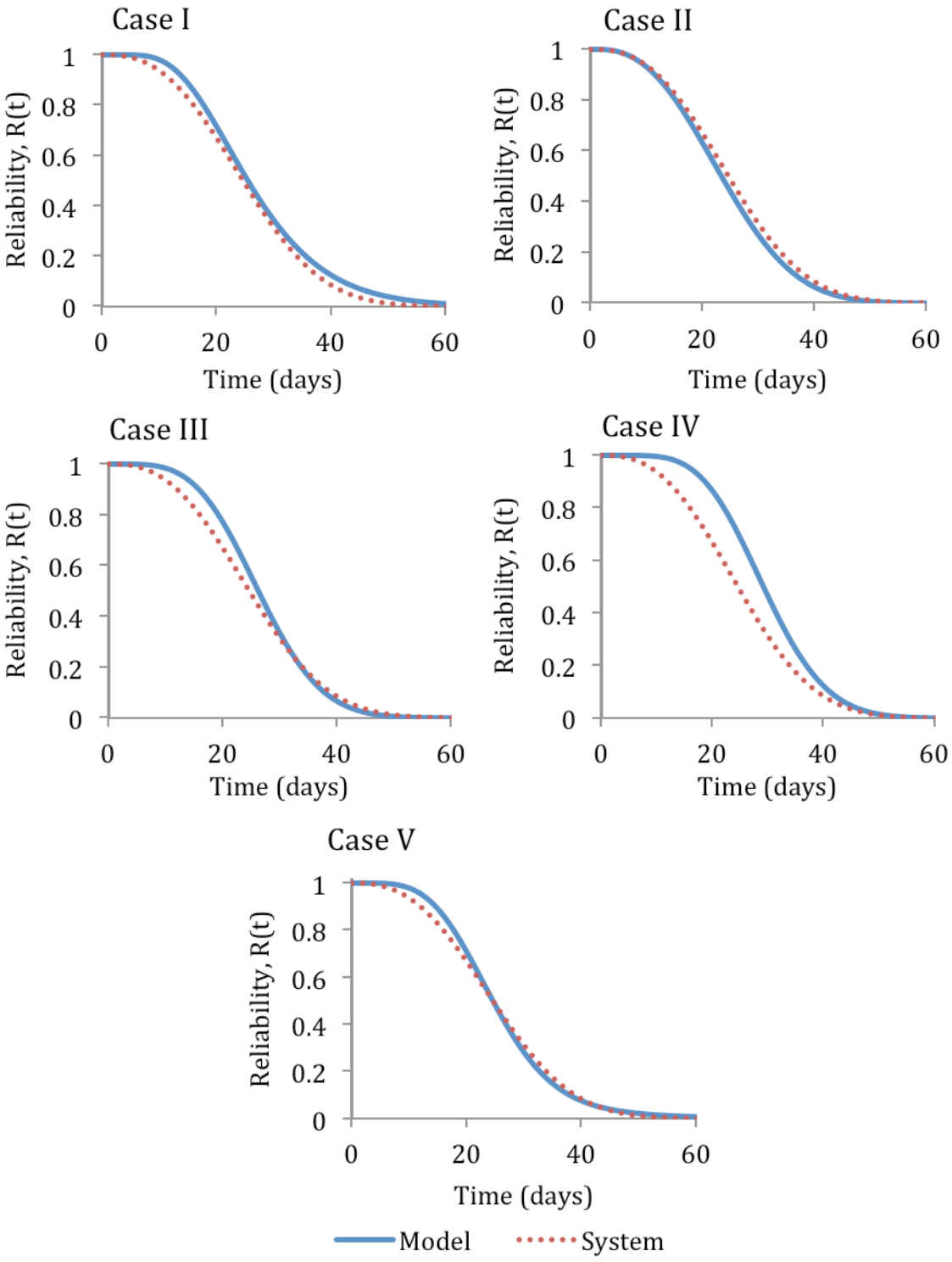



Figure 28. Comparison of model reliability functions and FP07 denitrifying biofilter system performance reliability function for the five component configurations analyzed.

# Chapter 4: Laboratory Biofilter Study

To allow more controlled study of these denitrification systems, laboratory-scale biofilters were designed and constructed. The use of laboratory-scale biofilters allows us to closely monitor microbial community structure and biofilter performance while controlling many environmental and operational parameters such as hydraulic retention time, nitrate concentration, and water depth. The laboratory biofilters provide a platform upon which experiments can be designed to test many hypotheses. In this thesis, the design and startup of the laboratory biofilters are presented, along with a change in biofilter operation. The goal of this change was to obtain different levels of diversity in the two biofilters to test the hypothesis that a less diverse community will result in less stable biofilter performance.

## 4.1 Methodology

### 4.1.1 Reactors

The laboratory-scale biofilters consist of 13cm x 46cm x 30.5cm reactors made of 0.236" thick clear extruded acrylic. Acrylic sheets were scored and snapped to size using an acrylic cutting tool. Joints in the acrylic sheets were adhered using Weld-On® 3 Solvent Cement (IPS Corporation, Compton, CA) and sealed with 100% silicone sealant (Dow Corning, Midland, MI).

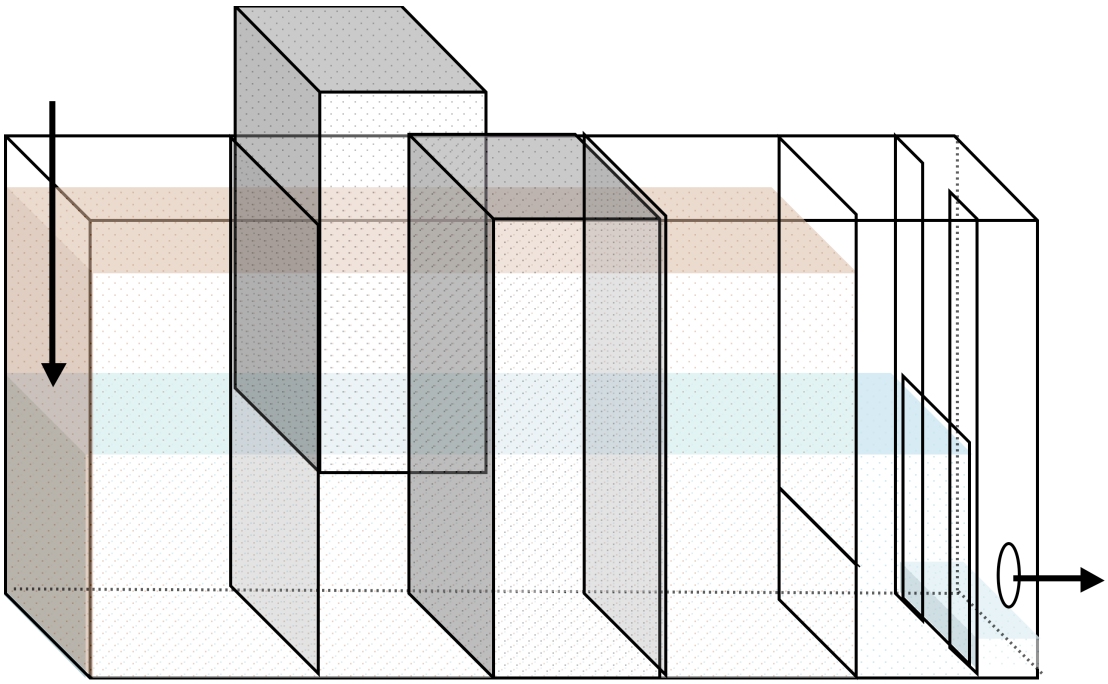


Figure 29. Diagram of laboratory-scale denitrifying woodchip biofilter. Brown area illustrate woodchip bed section and arrows indicate the direction of flow. Removable aluminum mesh sampling ports were inserted at full and half-depth of the biofilter to allow for woodchip samples to be taken. An adjustable overflow weir controlled water level in the biofilters.

The reactors are composed of two compartments: a large woodchip bed compartment and an overflow weir/drain compartment (Figure 29). The two compartments are separated by an acrylic baffle wall that extends from the height of the reactor to 5 cm from the bottom. The opening between the bottom of the baffle and the bottom of the reactor is screened with a ¼"aluminum mesh to prevent woodchips from leaving the woodchip bed compartment. An adjustable overflow weir consisting of a rubber-backed acrylic weir piece held tightly in place against a U-shaped frame with screws tightened through the drain side reactor wall controls water level in each biofilter. This allows for the water level in the reactor to be adjusted by exchanging weir pieces for others of different size.

Two woodchip sampling ports were constructed and inserted into the woodchip section of each biofilter. These ports consist of 13cm x 9cm x 34cm rectangular prisms of ¼" aluminum mesh with a removable mesh bottom. This allows for woodchip samples to be taken from the bottom of the sampling port, and woodchips above them to occupy the space that has been vacated by removal of the woodchip sample. Each biofilter was installed with two of these sampling ports, one that was inserted all the way to the bottom of the reactor and one that was inserted halfway into the reactor, with the space below occupied by woodchips. This allows sampling of woodchips from both the full depth and half depth of the reactor. The spaces where these ports are inserted are separated from the rest of the woodchip bed with ¼" aluminum mesh to prevent the woodchips from collapsing in the space vacated when the sampling port is removed.

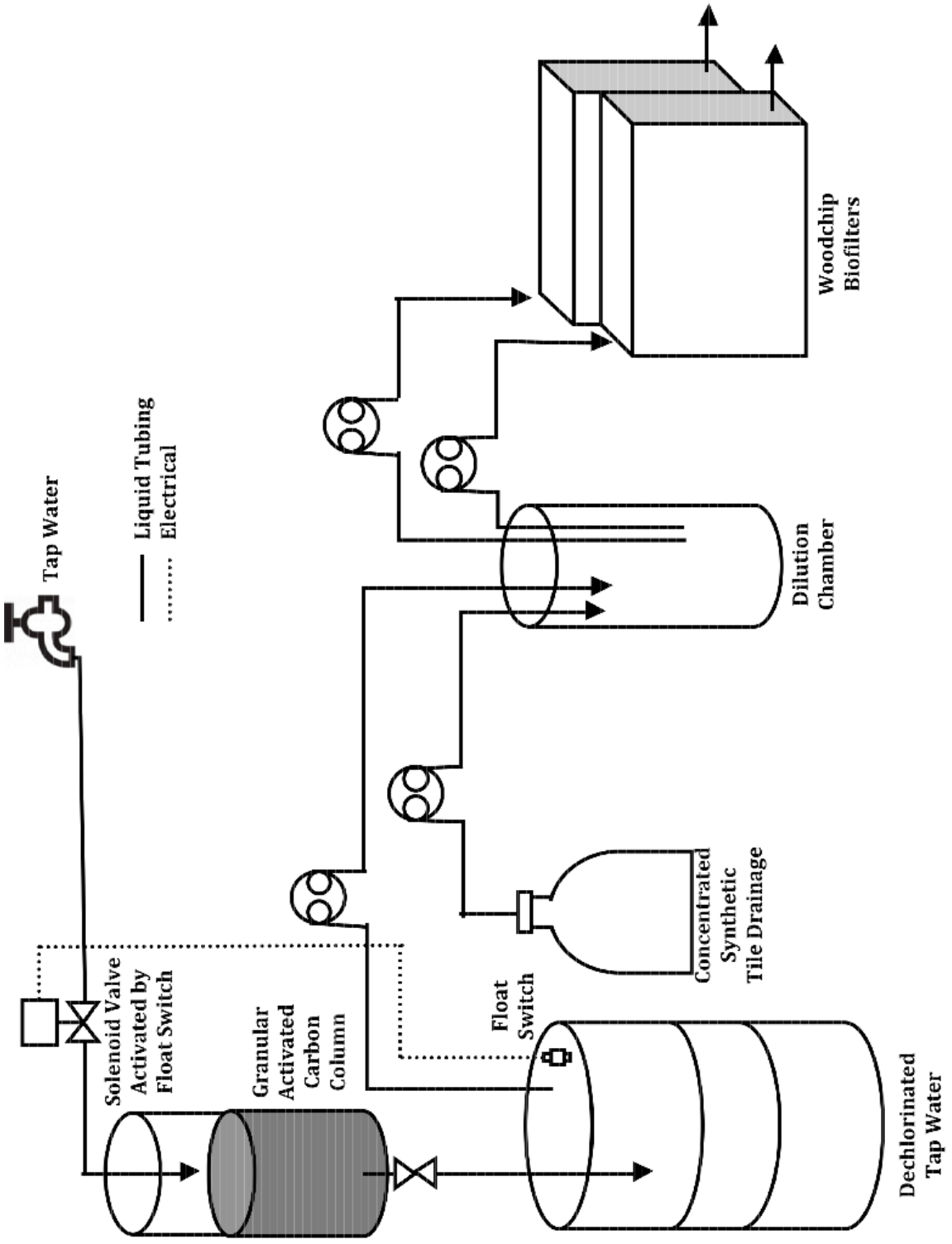



Figure 30. Laboratory denitrifying biofilter setup.

### 4.1.2 Reactor Setup

After some refining during the startup period, the final reactor setup includes several components (Figure 30). A 55-gallon drum was used to store dilution water. This drum was maintained at capacity with a float switch connected to a solenoid in the hose from the faucet. When the solenoid valve opened, the water flowed through a granular activated carbon column to remove chlorine and into the 55-gallon drum. This water was then pumped to a dilution chamber in which it was mixed with concentrated synthetic tile drainage. This diluted synthetic tile drainage was then pumped to the denitrifying bioifilters. In this section I will describe each of these components in detail with the exception of the synthetic tile drainage, which will be discussed in Section 4.1.3.

Peristaltic pumps consisting of Masterflex® L/S 7523-60 drives equipped with Model 77200-60 pump heads were used for all pumping requirements in the laboratory biofilter setup (Cole-Parmer, Vernon Hills, IL). The pump from the synthetic tile drainage feed bottle to the mixing chamber was equipped with Masterflex® L/S 13 Norprene tubing. The pump from the water reservoir to the mixing chamber and the pumps from the mixing chamber to each biofilter inlet were equipped with Masterflex® L/S 14 Norprene tubing.

Municipal tap water was used as dilution water for the synthetic tile drainage biofilter feeds. This water was first dechlorinated using a granular activated carbon (GAC) column. The GAC column consisted of a 52 cm tall x 21 cm diameter acrylic cylinder with 35 cm of bed depth. At the overflow head height of 52 cm, this corresponds to a throughput of approximately 0.135 liters/min and an empty bed contact time of 90 minutes. This water then flowed into a stainless steel 55-gallon drum reservoir. Using the biofilter flow rate ranges in this study, the retention time of water in the reservoir ranged from 8 to 12 days,

allowing it to reach room temperature before being pumped into the biofilters. Water level in the reservoir was maintained at approximately 55 gallons with a Madison M8000 float switch (Madison Company, Branford, CT) connected to an ASCO® Red-Hat® II solenoid valve (ASCO, Florham Park, NJ) in the water line from the faucet to the GAC column. When the water level in the reservoir drops 2", the solenoid opens and water begins flowing through the GAC column into the reservoir.

In the initial reactor setup, the tubing from the dilution water pump and concentrated synthetic tile drainage feed pump came together with a Y-connector and mixing occurred in-line. Because of inconsistent influent concentrations, the reactor setup was changed on 3/4/11 to include a separate mixing chamber to dilute the synthetic tile drainage. This mixing chamber consisted of a 39 cm tall by 18.5 cm diameter acrylic cylinder that was continuously mixed with a stir bar. Diluted synthetic tile drainage was then pumped from the same mixing chamber to each biofilter, keeping consistency in the influent concentrations seen by the two biofilters.

### 4.1.3 Operation

Woodchips used in the laboratory biofilters were collected from the FP07 field biofilter in Decatur, Illinois. Woodchips collected from a stockpile above the biofilter and extracted from the bottom of the biofilter were mixed together at an approximate ratio of 1:1. Porosity values of the woodchip beds were determined at the time of inoculation by filling each biofilter with woodchips and measuring the weight before and after filling the woodchip bed with water.

A concentrated synthetic tile drainage feed recipe (Table 13) was used and diluted before being pumped into the biofilters to allow for less frequent feed bottle changes. This

synthetic tile drainage used in the biofilters was derived from a basic bacterial and fungal media (Tanner 1997) that was scaled to be similar in chemical composition with respect actual tile drainage major mineral components (Table 16). The nitrate concentration of the synthetic tile drainage was maintained at 15 mg/L $NO_3$-N, comparable to what is seen in subsurface agricultural runoff in Central Illinois (Kalita et al. 2006). For the first two weeks of flow-through operation, the yeast extract concentration was 40 times larger than that of the final recipe in Table 13. This was scaled back to decrease the carbon content of the reactor feed. This synthetic tile drainage was made by combining the potassium nitrate, yeast extract, and mineral solution according to the recipe in Table 13 and autoclave sterilizing for 60 minutes. Once the liquid cooled down to room temperature, the metal solution was filter sterilized and added to the synthetic tile drainage.

Table 13. 10X concentrated synthetic tile drainage medium recipe derived from a basic bacterial and fungal media (Tanner 1997) that was scaled to be similar in chemical composition with respect to major mineral components as actual tile drainage (Blowes et al. 1994; Stone and Krishnappan 1997).

| 10X Synthetic Tile Drainage Medium | | |
|---|---|---|
| $KNO_3$ | 1.08 | g/L |
| Mineral Solution | 2.5 | mL/L |
| Trace Metal Solution | 0.625 | mL/L |
| Yeast Extract | 0.025 | g/L |

Table 14. Trace metal solution recipe for use in synthetic tile drainage derived from a basic bacterial and fungal media (Tanner 1997).

| Trace Metal Solution | | |
|---|---|---|
| Nitrilotriacetic acid | 2.0 | g/L |
| Adjust pH to 6 with KOH | | |
| $MnSO_4 \cdot H_2O$ | 1.0 | g/L |
| $Fe(NH_4)2(SO_4)2 \cdot 6H_2O$ | 0.8 | g/L |
| $CoCl_2 \cdot 6H_2O$ | 0.2 | g/L |
| $ZnSO_4 \cdot 7H_2O$ | 0.2 | g/L |
| $CuCl_2 \cdot 2H_2O$ | 0.02 | g/L |
| $NiCl_2 \cdot 6H_2O$ | 0.02 | g/L |
| $Na_2MoO_4 \cdot 2H_2O$ | 0.02 | g/L |
| $Na_2SeO_4$ | 0.02 | g/L |
| $Na_2WO_4$ | 0.02 | g/L |

Table 15. Mineral solution recipe for use in synthetic tile drainage derived from a basic bacterial and fungal media (Tanner 1997).

| Mineral solution | | |
|---|---|---|
| NaCl | 80 | g/L |
| KCl | 10 | g/L |
| $KH_2PO_4$ | 10 | g/L |
| $MgSO_4 \cdot 7H_2O$ | 20 | g/L |
| $CaCl_2 \cdot 2H_2O$ | 4 | g/L |

Table 16. Synthetic tile drainage composition compared to actual tile drainage composition. The final feed composition was calculated from the concentrated feed composition and the estimated composition of the dechlorinated tap water. This estimate was based on a study on groundwater quality in Champaign County (Sanderson and Zewde 1976). For constituents not included in this study (indicated with an asterisk), the final feed composition is based on a water concentration of 0 mg/L. All concentrations are in mg/L.

| | | | | Reference Tile Drainage | |
|---|---|---|---|---|---|
| | 10X Feed | Estimated Dilution Water | Final Feed | (1) | (2) |
| $NO_3$ | 662.31 | 0.47 | 66.65 | 15.0 | 44.20 |
| Ca | 2.72 | 53.13 | 48.09 | 94.80 | 106.00 |
| Mg | 4.93 | 29.51 | 27.06 | 22.70 | 21.40 |
| Na | 78.71 | 26.80 | 31.99 | 6.70 | 7.40 |
| K | 451.09 | | 45.11* | 1.60 | 0.73 |
| Si | 0.00 | 6.70 | 6.03 | | 4.98 |
| Cl | 138.06 | 1.81 | 15.43 | 21.20 | 12.30 |
| $SO_4$ | 20.14 | 0.38 | 2.36 | 38.00 | 20.90 |
| $NH_4$-N | 0.0459 | 0.8900 | 0.8056 | <0.05 | |
| $PO_4$-P | 5.6901 | | 0.5690* | <0.01 | |
| Fe | 0.0712 | 0.8800 | 0.7991 | <0.02 | |
| Mn | 0.2032 | 0.0221 | 0.0402 | 0.01 | |
| Cu | 0.0047 | | 0.0005* | <0.01 | |
| Zn | 0.0284 | | 0.0028* | <0.01 | |
| Co | 0.0310 | | 0.0031* | | |
| Ni | 0.0031 | | 0.0003* | | |
| Mo | 0.0050 | | 0.0005* | | |
| Se | 0.0052 | | 0.0005* | | |
| W | 0.0078 | | 0.0008* | | |

(1) Blowes et al. 1994
(2) Stone and Krishnappan 1997

The initial water levels in the R1 and R2 biofilters were 20.5 and 19 cm, respectively. Based on reactor geometry and estimated woodchip bed void volume, these water levels correspond to a 5.76 L woodchip bed void volume for both reactors. During the first two weeks, synthetic tile drainage was recirculated through the biofilter to ensure denitrifying microbes would not be washed out of the system. Flow-through operation of the biofilters began after this two-week seeding period. Continuous flow through the

biofilters remained for the rest of the study period in this thesis. The initial theoretical woodchip bed retention time was 2 hours for both reactors. This was increased to 3 hours for 18 days, then again to 6 hours for 9 days, and finally to 12 hours at which it remained for the remainder of the study period. The water level in the R2 biofilter was maintained at 19 cm for the entire study period, while the water level in R1 was changed from 20.5 to 11 cm on 8/27/11.

### 4.1.4 Sampling

Influent and effluent samples from the biofilters were collected at least twice per week to quantify nitrate removal performance. Effluent samples consisted of a 200 mL composite sample collected from the drain line. Influent samples were collected from the influent tubing coming from the mixing chamber. Influent and effluent samples were filtered using 0.2 μm syringe filters (Cole-Parmer P/N EW-02915-04, Vernon Hills, IL) and stored in 10 mL polyvinyl vials (Dionex P/N 055058, Sunnyvale, CA) for Ion Chromatography $NO_3$-N analysis.

Fifteen gram woodchip samples were collected weekly from the bottom of each sampling port for microbial community analysis and processed immediately as previously described (Porter 2011). These woodchip samples were placed into autoclaved 250 mL Nalgene bottles (Thermo Fisher Scientific, Rochester, NY) and 110 mL of Ringer's solution (Oxoid, Cambridge, UK) was added. The bottles were then shaken overnight in a 30˚C temperature controlled room. The following day, the woodchip wash for each sample was centrifuged at 5,000 g for 3 min to concentrate the microorganisms into a pellet. 2.5 mL of 1X PBS (Fisher Bioreagents #BP665-1) and five autoclaved 5 mm glass beads were added to each pellet and the samples were vortexed at full speed for 2 min. Each sample was then

centrifuged for 5 min at 750 g and the supernatant was collected and stored at -20˚C for future microbial analysis. Excess woodchips from inoculation of the biofilters were added to the top of the sampling ports as needed to replace the woodchips used for samples.

### 4.1.5 Analysis

Beginning in July 2011, 10 gram woodchip samples were taken weekly for denitrifying enzyme assay (DEA) analysis based on the acetylene block method described by Tiedje and colleagues (1989). Each woodchip sample was placed into an acid-washed and autoclaved 125 mL Wheaton bottle (Wheaton Science Products, Millville, NJ). 75 ml of solution containing 15 mg/L $NO_3^-$-N and 0.1g/l chloramphenicol in Type 1 water was added. The addition of chloramphenicol inhibits bacterial protein synthesis, extending the period of linear reduction of nitrate to nitrous oxide from which a denitrification rate can be determined. The Wheaton bottles were capped with open top caps with chlorobutyl septa (Wheaton Science Products, Millville, NJ) and the bottle headspace was flushed with helium gas for 10 minutes. Finally, 15 ml of the headspace gas was removed and replaced with acetylene to give an approximate headspace acetylene concentration of 20%.

At each sampling time point, a 12 mL gas sample and 4 mL liquid sample were removed from each bottle. The gas samples were transferred to 10 mL BD Vacutainer® evacuated serum tubes (BD Medical, Franklin Lakes, NJ) for storage until nitrous oxide analysis. The liquid samples were filtered using 0.2 µm syringe filters (Cole-Parmer P/N EW-02915-04, Vernon Hills, IL) and stored in 10 mL polyvinyl vials (Dionex P/N 055058, Sunnyvale, CA) for Ion Chromatography $NO_3^-$-N analysis. 16 mL of 90:10 He:Acetylene mixture was then added to the headspace of each bottle to replace the sample volume

removed. After completion of the denitrification enzyme assay, the dry weights of the woodchip samples used were measured after drying in a 105° C oven for 48 hours.

Gas samples were analyzed for $N_2O$ using a Shimadzu GC-2014 gas chromatograph equipped with an AOC-5000 Auto Injector (Shimadzu Corporation, Nakagyo-ku, Kyoto, Japan). The injection volume for analysis was 3 mL and samples were diluted to result in final concentrations in the detection range of 0.1 to 7 ppm. Linear regression of standards was used to determine the $N_2O$ concentration of the samples.

During the startup period, liquid samples were analyzed for $NO_3$-N using Hach NitraVer5 Reagent Powder Pillows (Hach Product # 2106169, Hach Corporation, Loveland, CO) and a Hach DR-4000u Spectrophotometer (HACH method 8039). After 4/4/11, liquid samples were analyzed for $NO_3$ -N using an ICS-2000 ion chromatograph system equipped with an AS-40 automatic sampler and Chromeleon® Chromatography Management Software (Dionex, Sunnyvale, CA). The chromatograph setup included a 0.25 µl sampling loop and an IonPac® AS18 4x250 mm hydroxide-selective anion-exchange column maintained at 30°C. The ion chromatograph system was equipped with a SRS 300 Self-Regenerating electrolytic suppressor with the suppressor current set at 80mA. The eluent used with this column was a 32 mM solution of KOH. The flow rate was 1.0 mL $min^{-1}$. Linear regression of 100, 50, 25, 12.5 and 6.25 ppm $NO_3$ standards was used to determine the $NO_3$ -N concentration of the samples.

Total chlorine (free and combined) measurements in the dilution water were taken weekly using Hach DPD Free Chlorine Reagent Powder Pillows (Hach Product #2105569) and a Hach DR-4000u Spectrophotometer (HACH method 8021). No total chlorine measurement above 0.05 mg/L was ever found in the dilution water.

### 4.1.6 Tracer Study

A tracer test was conducted on 11/5/11 and 11/6/11 to determine the residence time distribution for each reactor. A step input of 100 ppm bromide was used as a tracer. Effluent samples were collected at intervals of 30 minutes or 1 hour for 36 hours. Samples were analyzed for bromide using the same ion chromatograph method presented earlier for nitrate analysis.

The effluent bromide concentrations were used to determine the value of the cumulative age distribution function, $F(t)$, at each time point (Equation 31).

$$F(t) = \frac{C_{out}(t)}{C_{in}} \quad (31)$$

The derivative of this function yields the exit age distribution, $E(t)$, the instantaneous fractional rate at which input tracer mass is leaving the reactor (Equation 32)

$$E(t) = \frac{dF(t)}{dt} \quad (32)$$

Equation 32 was approximated by Equation 33 using the discrete data obtained in the tracer study.

$$\left(\frac{dF}{dt}\right)_i \approx \left(\frac{\Delta F}{\Delta t}\right)_i = \frac{F_i - F_{i-1}}{t_i - t_{i-1}} \quad (33)$$

The average hydraulic retention time was then determined using the exit age distribution from the tracer data and Equation 34.

$$\bar{t} = \frac{\sum_{\text{all } i} t_{i,ave} E_{i,ave} \Delta t_{i,\text{int}}}{\sum_{\text{all } i} E_{i,ave} \Delta t_{i,\text{int}}} \quad (34)$$

## 4.2 Results

### 4.2.1 Reactor Performance

The woodchip bed porosities for the biofilters determined at the time of inoculation were 0.60 and 0.62 for R1 and R2, respectively. These measured porosities are similar to the 100% woodchip bed porosity of 0.57 determined by Wildman (Wildman 2002). The initial woodchip bed porosity for each biofilter is used for theoretical hydraulic retention time determination throughout the study.

After the two week inoculation period, flow-through operation of the biofilters began with a theoretical woodchip bed retention time of two hours. Shortly after the change to flow-through operation, the effluent became optically clear. This allowed for nitrate to be analyzed using the Hach Specrophotometer. These initial measurements showed nearly 100% nitrate removal in both biofilters. This removal rate is significantly higher than what has been seen in the field (Chun et al. 2010). We believed that this discrepancy was caused by the high carbon content of the initial media recipe. Therefore, the amount of yeast extract in the synthetic tile drainage was reduced by a factor of 40 on 12/2/2010. This resulted in an immediate decrease in removal rate, as seen in Figure 31.

After this change, the difference in nitrate concentrations between the influent and effluent was not great enough to use as a baseline for future experiments and reliability analysis. Therefore the theoretical woodchip bed retention time was increased to 3, 6, and then 12 hours by adjusting flow rate until there was adequate separation between the influent and effluent nitrate concentrations (Figure 31).

During the entire startup period, frequent changes in influent concentrations were seen. We had two hypotheses as to the cause of these changes. The first was that the in-line mixing was inadequate and we were not achieving consistent pumping. To address this concern, a continuously stirred dilution chamber was added to the reactor setup on 3/4/11. This resulted in less rapid changes in influent concentration, although there was still some variation. The second hypothesis of the cause of the inconsistent influent concentration was a concern with the accuracy of the measurement method. To address this concern, I started analyzing samples using ion chromatography as well as the Hach method. A comparison of the results from the two measurement methods demonstrated that a significant amount of the variation seen in the samples was likely caused by the inconsistency in the Hach measurement method (Figure 32). After seeing these results, ion chromatography was used to analyze samples for nitrate for the rest of this study.

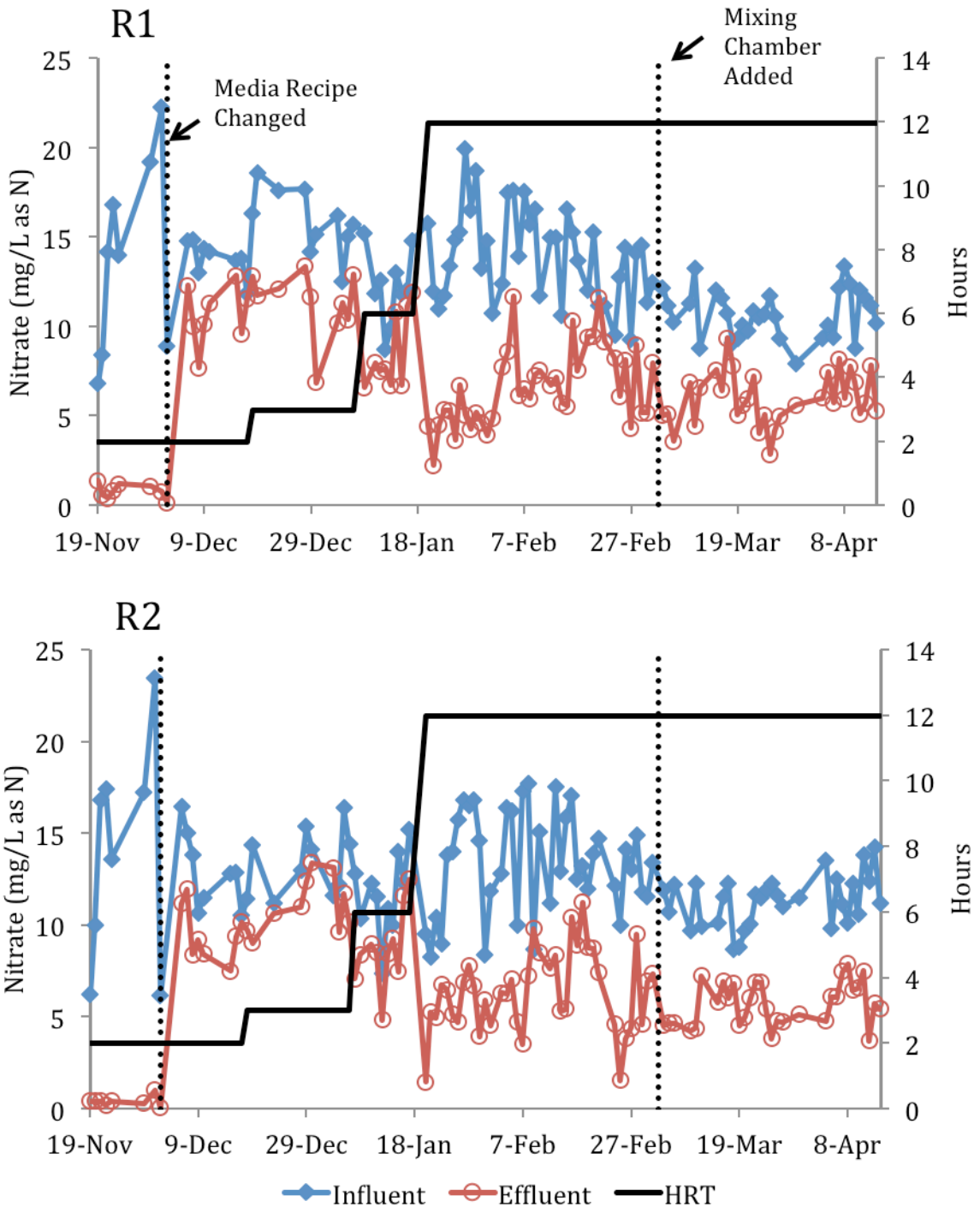



Figure 31. Laboratory scale denitrifying biofilter influent and effluent nitrate concentrations from 11/19/10 to 4/13/11 measured using the Hach NitraVer5 reagent. On 12/2/10 the amount of yeast extract was reduced by a factor of 40. On 3/4/11 a mixing chamber was added to the reactor setup. During this period the hydraulic retention time was increased until a significant separation between influent and effluent concentrations was seen.

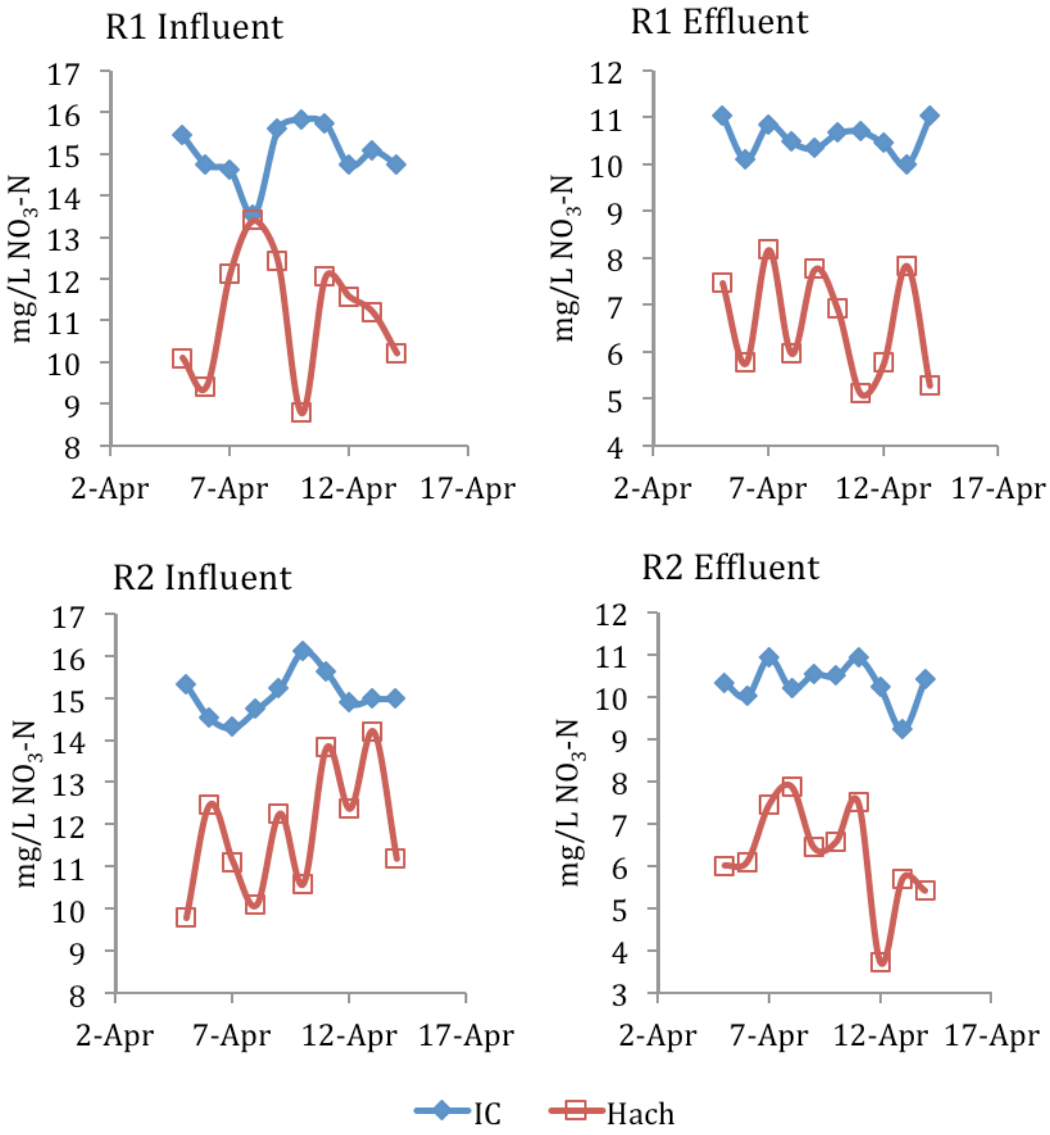



Figure 32. Comparison of influent and effluent nitrate concentrations measured using the Hach NitraVer5 reagent and ion chromatography. Inconsistency in the Hach measurements led to the use of ion chromatography for nitrate analysis for the remainder of the study.

After adding the mixing chamber and changing to ion chromatography for nitrate analysis, the biofilters were both operated at a 12 hour theoretical woodchip bed retention time for several months to allow for monitoring of the system performance stability. During this period, the biofilters exhibited very stable performance, showing

approximately 30% removal (Figure 33). One R1 effluent data point (7/7/11) that was significantly lower than its R2 counterpart resulted from a clog in the influent tubing, which increased the hydraulic retention time. Other than that date, the two biofilters exhibit very similar removal over the entire period.

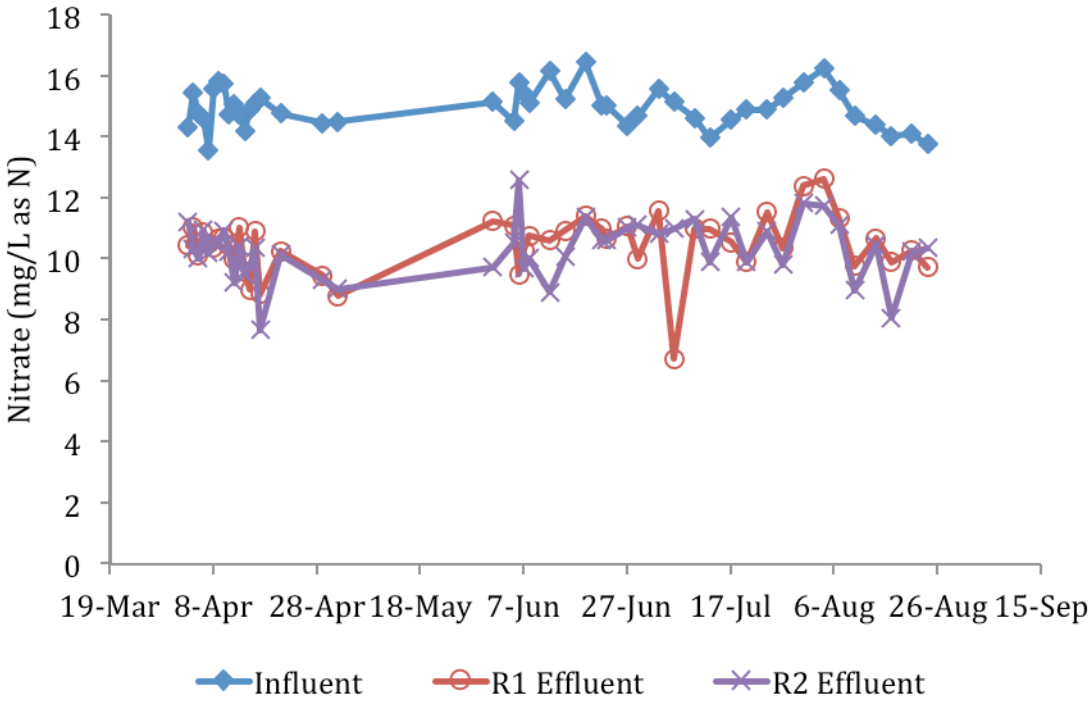



Figure 33. Laboratory scale denitrifying biofilter influent and effluent nitrate concentrations from 4/4/11 to 8/25/11 at a theoretical woodchip bed hydraulic retention time of 12 hours. During this period when both reactors were operated at approximately the same weir height and hydraulic residence time, nitrate removal in both biofilters was very stable at approximately 30%.

As discussed earlier, it is hypothesized that a varying environment will lead to a less diverse microbial community and less stable performance. With this in mind, after both biofilters demonstrated that their performance was stable, the water level in one reactor was lowered to about half of its previous height on 8/27/11. The theoretical hydraulic

residence time in the woodchip bed was held constant at 12 hours by adjusting the influent flow rate accordingly (Table 17).

Table 17. Laboratory biofilter operational parameters during the test period presented in this thesis.

| | R1 (4/4/11-8/27/11) | R1 (8/27/11-11/4/11) | R2 (4/4/11-11/4/11) |
|---|---|---|---|
| Weir height (cm) | 20.5 | 11 | 19 |
| Woodchip reactor volume (L) | 9.60 | 5.15 | 9.25 |
| Measured initial woodchip bed porosity | 0.60 | 0.60 | 0.62 |
| Estimated water volume in woodchip bed (L) | 5.76 | 3.09 | 5.76 |
| Influent flow rate (mL/min) | 8.0 | 4.3 | 8.0 |
| Theoretical woodchip bed HRT (hours) | 12 | 12 | 12 |

After the water level was lowered in the R1 biofilter, the reactor performance immediately improved (Figure 34). On 8/29/11, woodchips were added to the top of the biofilter sampling ports, resulting in wood 'dust' falling into the exposed water and increased nitrate removal in both reactors for that day. The following day, both reactors had returned to their level of performance seen immediately after the change in water level. While the nitrate removal efficiency in R2 remained at approximately 30% during this operational change, the removal efficiency in R1 jumped to around 55% and remained consistent at this level. This difference in reactor performance between the two reactors was somewhat unexpected, and therefore the weir height in R1 was kept at this lower level to explore possible explanations. For a three day period (10/18-10/20) in the biofilter

operation, the room temperature raised significantly (6-7 degrees C), resulting in high nitrate removal in both biofilters. Both before and after the change in operation, the nitrate load reduction was very similar between the two biofilters, even though the change resulted in an R2 reactor volume nearly twice that of R1 (Figure 35).

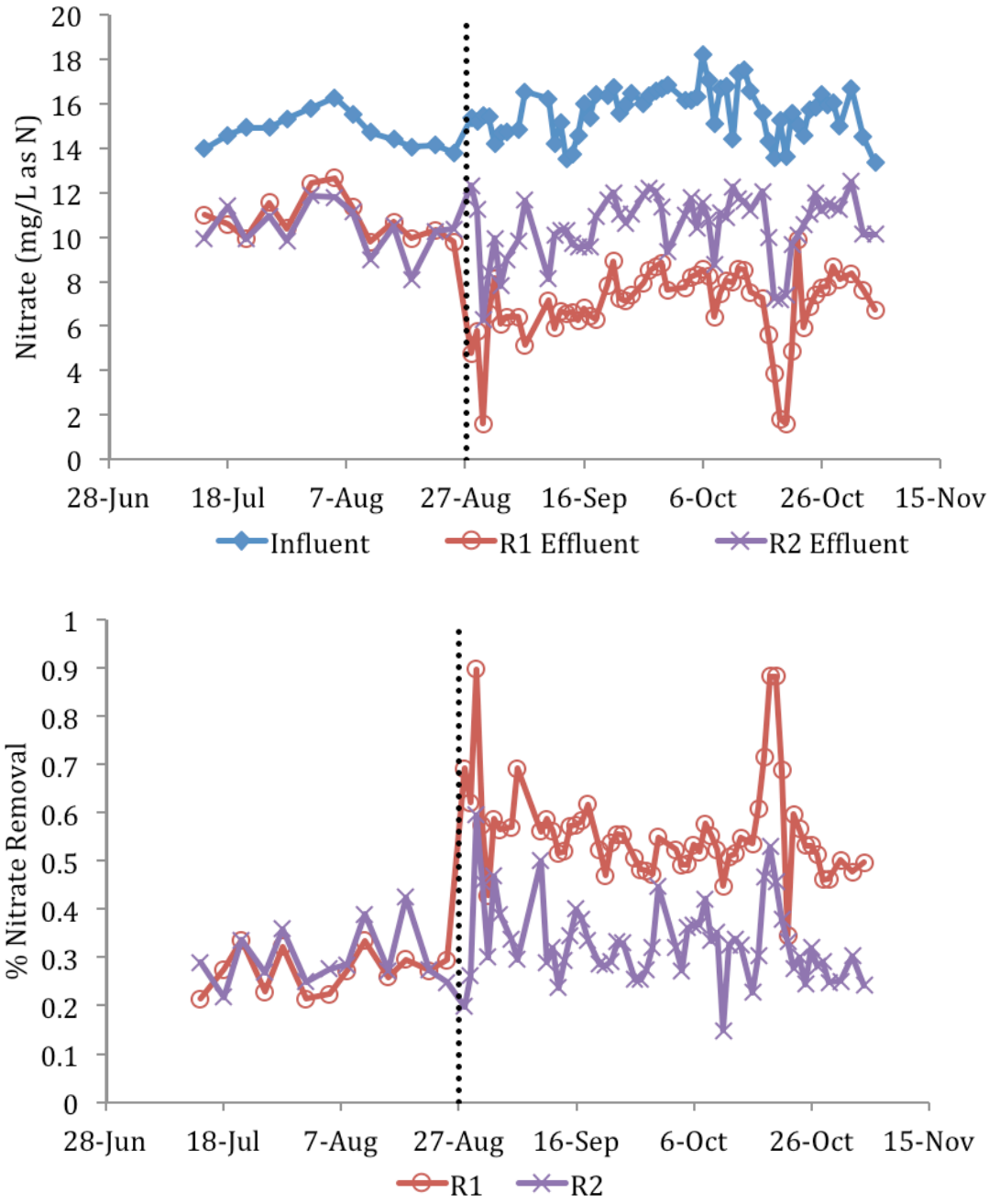



Figure 34. Laboratory scale denitrifying biofliter performance at a theoretical woodchip bed hydraulic retention time of 12 hours. The weir height in R1 was lowered on 8/27/11 (marked by the vertical line), at which time the system immediately showed improved nitrate removal. This improved nitrate removal continued through the remainder of the experiment. Woodchips were added to the top of the biofilter sampling ports on 8/29/11, resulting in wood 'dust' falling into the exposed water and increased nitrate removal for that day. Between 10/17/11 and 10/20/11 the normal laboratory room temperature of 20 degrees Celsius was elevated to 27 degrees Celsius.

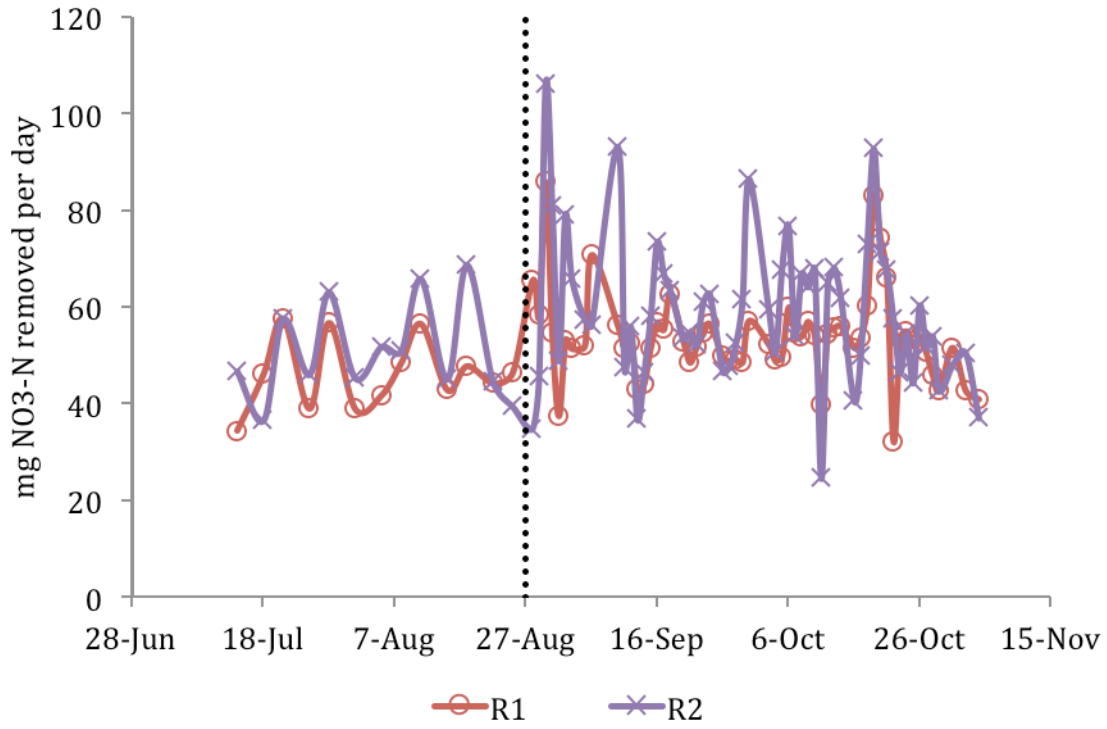



Figure 35. Laboratory scale denitrifying biofliter nitrate load removal calculated from influent and effluent nitrate concentrations and influent flow rate. The weir height in R1 was lowered on 8/27/11 (marked by the vertical line). While a difference in nitrate removal efficiency is seen in the two reactors, the total amount of nitrate being removed from in the system per day is similar.

### 4.2.2 Woodchip Moisture Content in Denitrifying Enzyme Assays

Moisture content measurements were taken for the woodchips that were used in the denitrifying enzyme assay, starting on the day of the change in operation of R1. All woodchip moisture content measurements fell between 0.55 and 0.85, with all of the submerged woodchips remaining above 0.70 (Figure 36). After 4 weeks of woodchip moisture content values similar to those of woodchips from the other ports, the woodchip moisture content in the R1 shallow port decreased until it reached around 0.55. The time at which the R1 shallow port reached this woodchip moisture content range coincides with

the trend reversal in the DEA results back to higher nitrate removal in the deep port. Another interesting trend is that the moisture content of the woodchips in the R2 deep port was consistently higher than that in the R2 shallow port, even though they were both completely submerged the entire period.

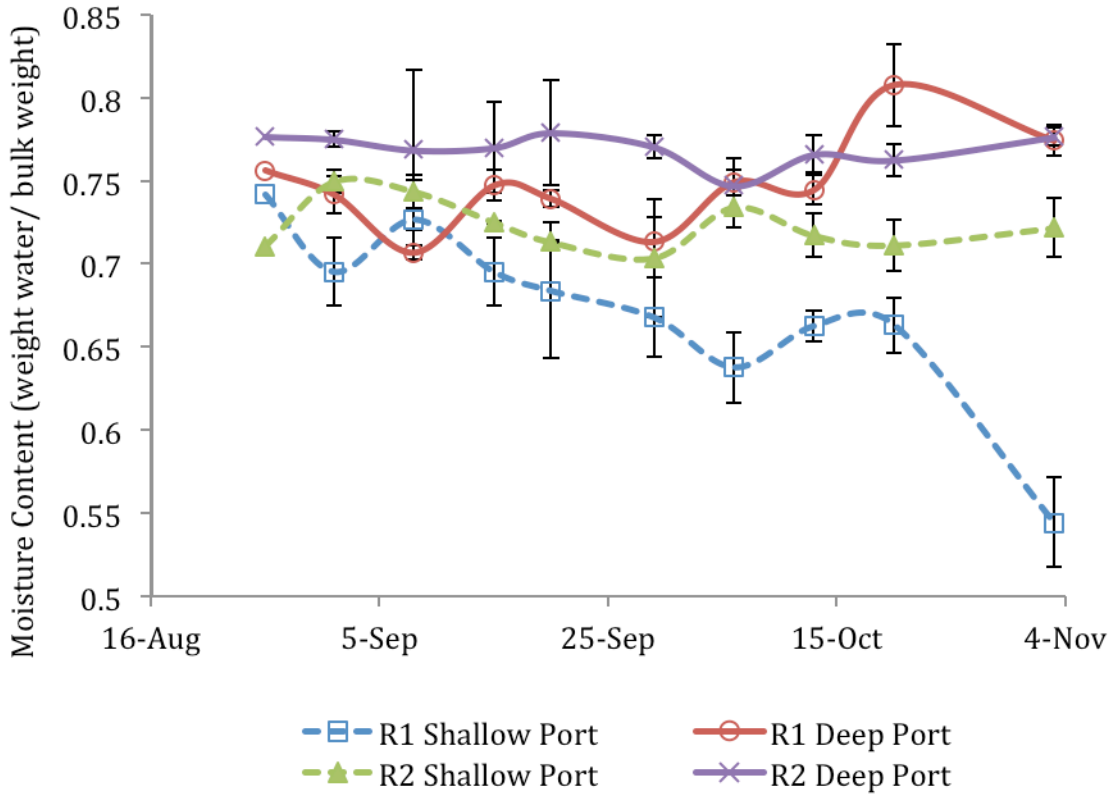



Figure 36. Woodchip moisture content values in denitrifying enzyme assays after change in water level. The woodchips from the R1 shallow port were no longer submerged after August 27 (first sample date), resulting in lower moisture content readings in the following weeks. Error bars indicate one standard deviation from the average value.

### 4.2.3 Tracer Study Results

A tracer study was performed on 11/5/11 and 11/6/11 to determine if a difference in actual hydraulic retention time could explain the difference in performance between the two biofilters after the change in operation. Measured influent and effluent bromide

concentrations were used to determine a cumulative age distribution (Figure 37) for each reactor using Equation32. These cumulative age distribution curves reveal that the mixing regime in the laboratory reactors falls between that of a CSTR and a PFR.

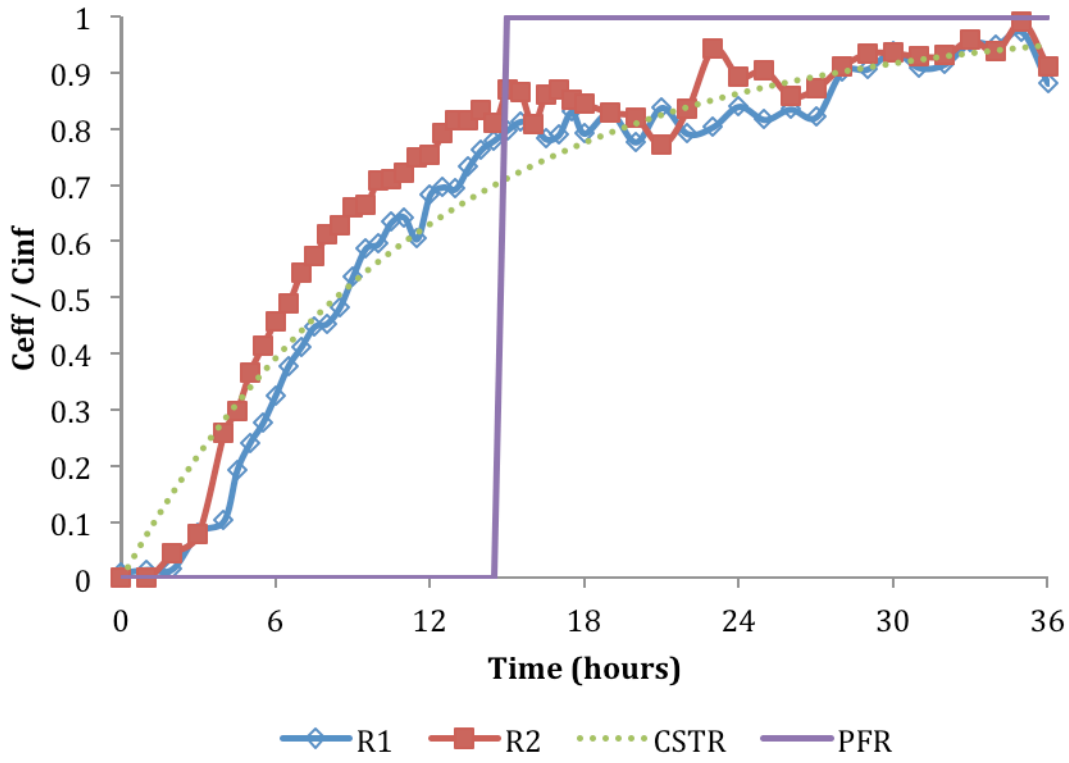



Figure 37. Laboratory biofilter step input tracer study cumulative age distribution curves were determined from influent and effluent bromide concentration measurements. Comparing experimental curves with those for ideal continuously stirred tank reactor (CSTR) and plug flow reactor (PFR) with 14.8 hour theoretical hydraulic retention times reveal that the mixing regime in the laboratory reactors falls between that of a CSTR and a PFR. Average hydraulic retention times were 8.84 hours for R1 and 7.52 hours for R2.

Average hydraulic retention times for both reactors were determined using this data and Equation 34. While both reactors had a theoretical hydraulic residence time of 12 hours in the woodchip bed, with 14.8 hours after including the overflow weir volume, the average hydraulic retention times (woodchip bed and overflow weir) determined from the tracer test were 8.84 hours for R1 and 7.52 hours for R2. This suggests that the actual

reactor volumes are not as large as calculated based on reactor geometry and initial woodchip bed porosity, either because of dead volumes in the reactor or a change in the woodchip bed porosity over time.

### 4.2.4 Denitrifying Enzyme Assay Results

Along with the traditional measure of denitrification using the acetylene block denitrifying enzyme assay method, which is nitrous oxide production, the modified assay method used in this thesis also provided data on the rate of nitrate removal. In a system designed to remove nitrate from tile drainage, the rate at which nitrate is removed in the system is relevant in understanding how the system is functioning, while nitrous oxide production measurements reveal the amount of nitrate removed through denitrification.

#### 4.2.4.1 Nitrate Removal in Denitrifying Enzyme Assays

Nitrate removal rates in the denitrifying enzyme assays averaged across replicates ranged from 0.00355 to 0.0212 mg N per dry gram of woodchip per hour (Figure 38). A large variation is seen in the nitrate removal rates, with relative standard deviation values up to 62%. This is not surprising when using a 10 gram sample of an inhomogeneous material such as woodchips. Comparing the nitrate removal rates between the woodchips in the deep and shallow ports from R2, more removal occurred in the deep port in almost every assay. The results from R1, which had the change in weir height, do not show the same trend. Initially, woodchips from the deep ports exhibited more nitrate removal, similar to what was seen in R2. However, for weeks 3-5 the nitrate removal was higher in the shallow ports, which contained woodchips from the unsaturated zone. The difference between the removal rates in the shallow and deep ports in this time period results from

both an increase in the rates in the shallow ports and a decrease in the rates in the deep ports. After the fifth week of operation at a lower water level, this trend reversed and the woodchips in the deeper ports once again exhibited more nitrate removal.

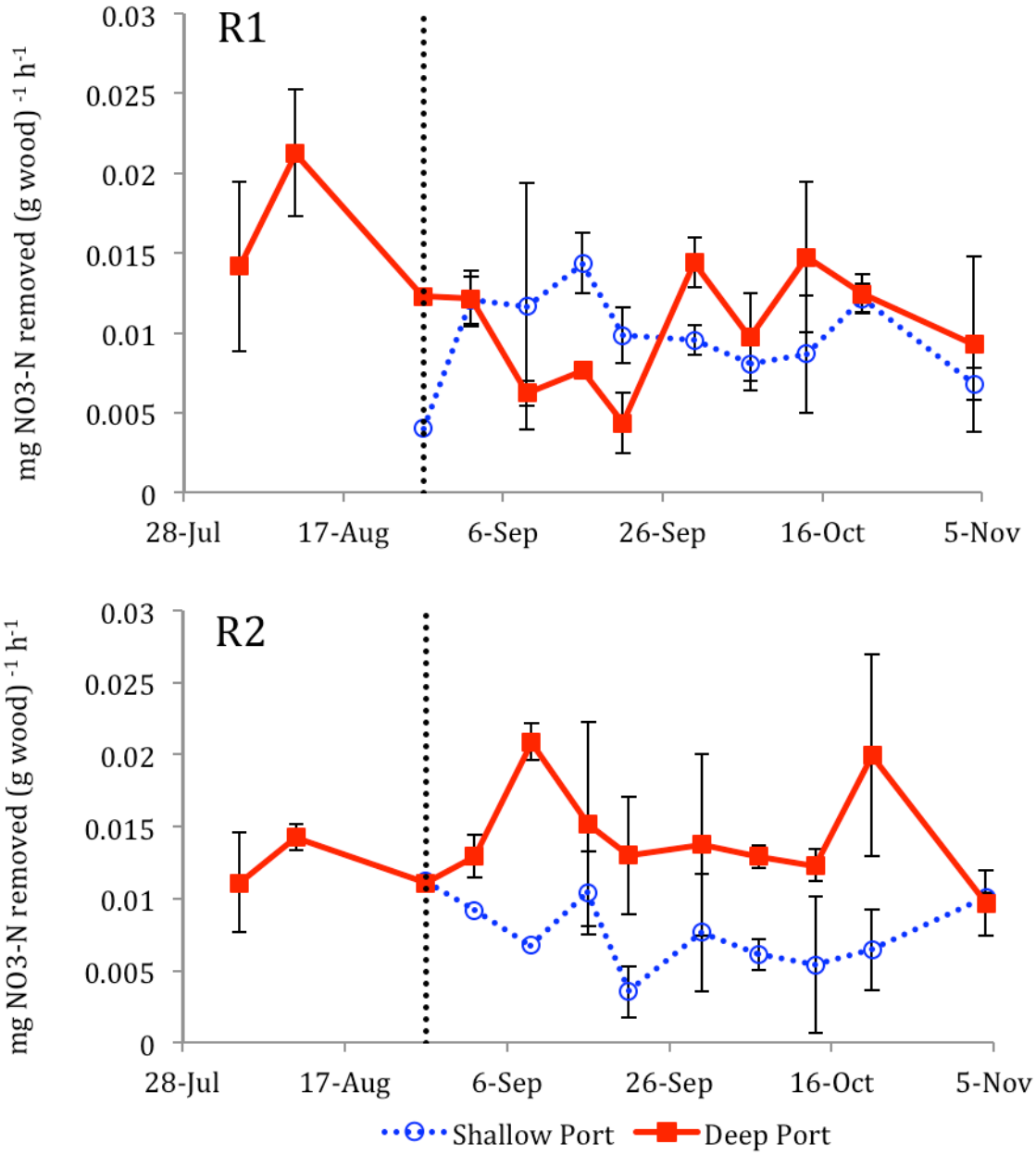



Figure 38. Denitrifying enzyme assay results showing mean nitrate removal rates per dry gram of woodchip from two replicates. Samples were taken approximately weekly from both the shallow and deep port of each laboratory biofilter after the water level in R1 was lowered on 8/27/11 (marked by vertical line) to monitor changes in microbial denitrification capabilities. Error bars indicate one standard deviation from mean value.

#### 4.2.4.2 Nitrous Oxide Production in Denitrifying Enzyme Assays

Nitrous oxide production measurements in the denitrifying enzyme assays reveal the denitrifying potential of the microbial community in the biofilters. Although gas samples were taken for every denitrifying enzyme assay, only 5 weeks of gas samples were analyzed because of chromatograph availability. Average nitrous oxide production rates across replicates ranged from 0.0003 to 0.00235 mg N per dry gram of woodchip per hour (Figure 39). A large variation was also seen in the nitrous oxide production rates, with relative standard deviation values up to 51%.

Comparing the nitrous oxide production rates in the deep and shallow ports from R2, a greater mean rate occurred in every assay (Figure 39), agreeing with the nitrate measurements (Figure 38). This consistency is not seen in the results from R1 where nitrous oxide production rates in the deep port are equal to or greater than those in the shallow port in every assay analyzed (Figure 39), while nitrate removal rates were greater in the shallow ports in some of these samples (Figure 38).

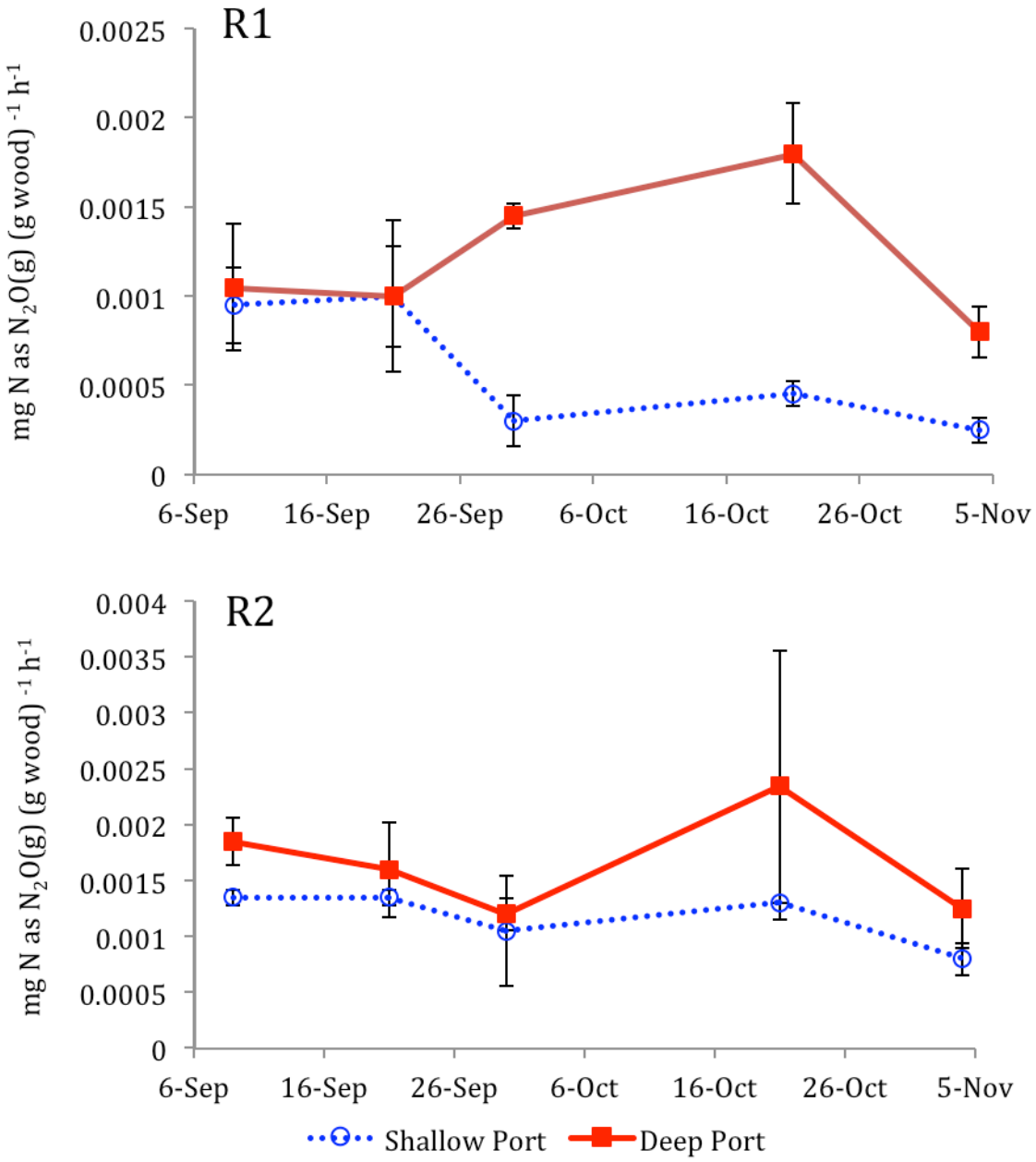



Figure 39.Denitrifying enzyme assay results showing mean nitrous oxide production rates per dry gram of woodchip from two samples. Samples were taken approximately weekly from both the shallow and deep port of each laboratory biofilter to monitor changes in microbial denitrification capabilities. Error bars indicate one standard deviation from mean value.

#### 4.2.4.3 Comparison of Nitrate Removal and Nitrous Oxide Production

Since the modified denitrifying enzyme assay method used in this thesis provides both nitrate removal and nitrous oxide production rates, a comparison between the two can be made. A significant discrepancy exists between the amount of nitrate-N removed and the amount of nitrous oxide-N produced in the denitrifying enzyme assays, with the nitrous oxide production being consistently around one-tenth the nitrate removal (Figure 40). Additionally, when comparing across ports, the trends between two rates are not always the same. For example, the R2 shallow port has a much higher mean nitrous oxide production rate than the R1 shallow port, but the R1 shallow port has a higher mean nitrate removal rate (Figure 40). Therefore, caution must be taken when comparing rates across ports based on only one of these rates. The reason for the discrepancy in nitrate removal and nitrous oxide production rates in the denitrifying enzyme assays in currently unexplained, and is discussed in more detail in Chapter 5.

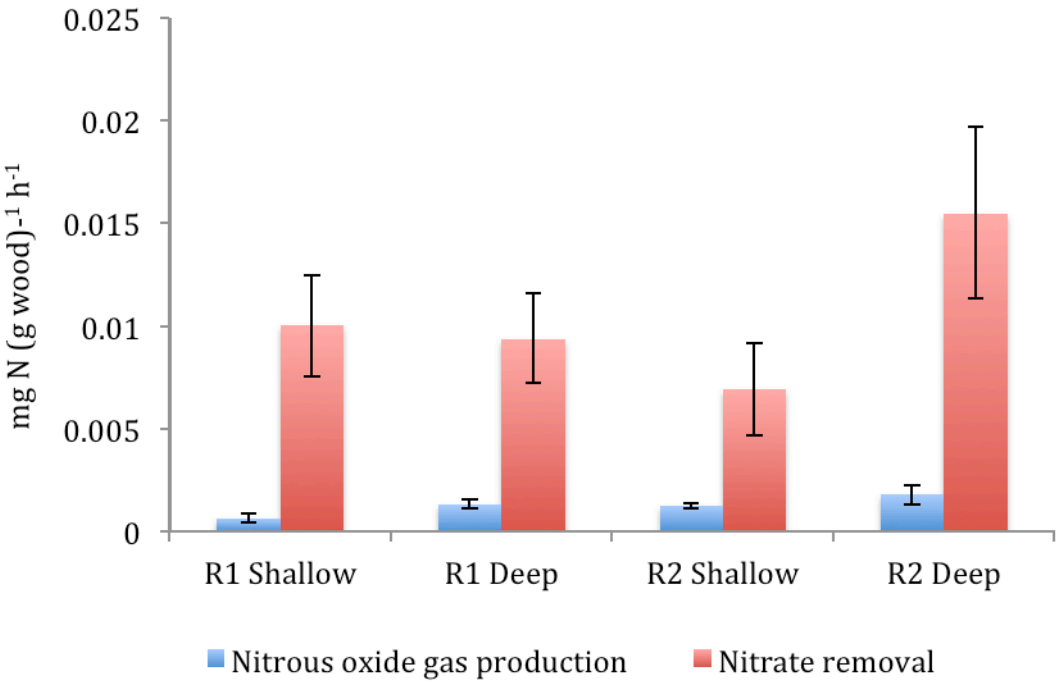



Figure 40. Comparison of nitrate removal and nitrous oxide production rates in denitrifying enzyme assays. Rates calculated from average rate in assays from dates that had both nitrous oxide and nitrate samples analyzed. Error bars indicate one standard deviation from mean value.

# Chapter 5: Discussion

In this work I have developed a method and associated software, the Ecosystem Reliability Analysis Tool (EcoReliAnT), to apply reliability engineering methods to engineered ecosystems. I applied this software to quantify the reliability of microbial populations and to model the reliability of engineered ecosystems as a configuration of functional microbial components in two engineered ecosystems. Additionally, I have established laboratory scale denitrifying biofilters that will allow study of denitrification systems with more control over system operation and closer monitoring of system performance and microbial community structure. This chapter consists of discussion of my research in four separate sections. First I will discuss my experience applying reliability techniques to engineered ecosystems in general. The next two sections will discuss the specific results of the reliability analysis of the external and field biofilter datasets, respectively. Finally I will discuss the laboratory denitrifying biofilters.

## 5.1 Application of Reliability Theory to Engineered Ecosystems

To my knowledge, the work in this thesis is the first time reliability methods have been applied to engineered ecosystems to explore ecosystem stability concepts. While some of the reliability theory methods used to analyze engineered ecosystems in this thesis are well established (i.e. life testing, parameter estimation using maximum likelihood estimation, chi-square goodness-of-fit testing, *k*-out-of-*n* system reliability determination), our method of determining reliability functions for microbial populations based on their population dynamics and modeling ecosystem performance using microbial functional

components is a completely new approach. In this section I will first discuss some challenges I encountered that are common in reliability engineering and how my method was adjusted to overcome them. I will then discuss challenges unique to the application of reliability methods to engineered ecosystems, how they were handled in this thesis, how they affect the reliability analysis, and how the methods can be improved to result in a more accurate reliability model.

### 5.1.1 Challenges Common in Reliability Analysis

There are challenges that I encountered when applying reliability analysis methods to engineered ecosystems that are commonly seen in reliability engineering. While the challenges may not be unique to this application, the solutions were developed specifically for the analysis in this thesis. In this section, I will discuss the challenges and effects of using small sample sizes and defining failure using a threshold.

#### 5.1.1.1 Sample Size and Reliability Analysis

As with most statistical methods, larger sample sizes yield more confidence in the results. Although both of the datasets analyzed in this thesis had quantifiable failures, the number of failures was much smaller than what would ideally be desired for reliability analysis. This situation is likely to be true for the reliability analysis of engineered ecosystems in general, since the method obtains all of the 'times to failure' from a single time series. The small sample sizes in this thesis affected my analysis in three ways: requiring me to group populations, consider censored data to be complete, and use the Weibull reliability model for all populations.

The small number of failures for each microbial population led me to group similar populations together to pool their 'times to failure' and form a larger sample size from which to estimate reliability function parameters. To form groups, I compared each microbial population to the system performance data, creating one group of populations that were present at higher abundance values during periods of good performance and another group that were present at higher abundance values during periods of poor performance.

For the more complex biofilter dataset, I used the Local Similarity Analysis technique (Ruan et al. 2006) to compare microbial population dynamics to system performance. One thing that is not captured in the LSA method when correlating individual populations to system performance is how combinations of the abundance values of microbial populations would compare to system performance. For example, if two microbial populations were truly redundant in function, it is possible that neither of their population dynamics would correlate to system performance individually, but the combination of the two would. This is because the system can perform its function with one population absent, as long as the other is present. After my experience analyzing the datasets in this thesis, I feel that while the LSA method was useful in my reliability analysis, there may be more comprehensive methods for grouping and incorporating microbial populations into the reliability model system.

The small sample sizes of failures for both microbial populations and system performance data also required me to ignore right-censoring that occurred in the determination of 'times to failure' in both analyses. Generally, accounting for censoring would be advised, as it yields reliability functions more accurately describing the failure

distributions seen in the system. However, if all of the failure data is censored, it is impossible to determine a reliability function. Because censoring was ignored, my analysis will underestimate the predicted 'times to failure' of the system itself and the microbial populations in the system.

Finally, the small sample size led me to use only the Weibull reliability model throughout the thesis. Although the EcoReliAnT software has a built-in goodness-of-fit test that automatically applies the chi-square test to the failure data and chosen reliability model, caution must be taken when using the chi-square test with small sample sizes (bins having expected frequencies less than 5) (Snedecor and Cochran 1989). Due to the small sample sizes analyzed in this thesis, only the Weibull model was used because of its versatility and common usage in reliability theory.

Collecting more failure data may prove difficult in these systems, as techniques commonly used in reliability engineering do not transfer directly to biological systems such as the denitrifying biofilters. When testing a manufactured product, an accelerated life test could be used, yielding more failure data and a more accurate reliability model. This type of test may not be possible when microbial population dynamics are being used to quantify reliability, since you cannot uniformly 'speed up' factors such as the growth rate of microbes or the time required to observe changes in microbial community structure. Additionally, changing environmental conditions in laboratory biofilters to 'accelerate' failure of the system would likely result in a failure directly caused by the environmental conditions. One alternative would be to combine data from several biofilters. This would not only provide more data and a more accurate model, but it would also yield a model that is more comprehensive rather than potentially site-specific.

#### 5.1.1.2 Failure Threshold Determination

The most important decision in the determination of a reliability function is how a 'failure' is defined through the use of a threshold value. In this thesis, defining thresholds at which failure of microbial populations occurred was somewhat of an art. To apply the reliability theory methods, failure must be observed. Tolerances and operational specifications are commonly used in engineering as failure thresholds. A system or component may be working, but if it is performing its function outside of the allowed tolerance range, it is considered to have failed.

Our case is a little different, since failure thresholds are used for both the system and its components. In engineered systems, component failure thresholds can be set at values known to result in system failure. For example, if the torque produced by a given motor is less than 80 ft-lbf the pump it is connected to cannot work, so a failure threshold of 80 ft-lbf is justified. In our case it is impossible to determine the abundance value for each population that is required for the system to be operational. However, since the microbial populations analyzed were only those correlated to system performance, using failure thresholds that are the same relative to each component's own time series data (e.g. average minus one standard deviation) in essence defines component failure based on system failure. For example, if the failure threshold of the average minus one standard deviation is used for both performance data and microbial abundance, the failures of populations and the system correspond to one another since the dynamics of the two datasets are correlated. The use of a threshold based on the average and standard deviation abundance or performance value also assured that failures were observed.

The choice of a threshold value can have a large impact on the estimated reliability function parameters. Because of this, further development of methods for defining threshold values and determining the sensitivity of reliability functions to changing threshold values would be worthwhile.

### 5.1.2 Challenges Unique to Engineered Ecosystem Reliability Analysis

In addition to the common challenges that I encountered when applying reliability analysis methods, there were several challenges that were unique to this application. In this section, I will discuss these unique challenges, how they were handled, and their effect on the analysis. I will also provide suggestions for possible reliability modeling improvements based on my experiences. The unique challenges in this section are presented in my opinion of their order of importance: (1) environmental and operational influences on system performance; (2) use of relative abundance data; (3) unknown system component configuration; and (4) microbial population 'life testing'.

#### 5.1.2.1 Environmental and Operational Influences on System Performance

The effect of outside influences on system performance is the largest challenge in the current approach of modeling the performance reliability of engineered ecosystems. In an engineered ecosystem, the system performance is affected by both the microbial community composition and the environmental and operational conditions. In this thesis, the system performance is modeled based on the microbial community only. This is appropriate for the external dataset analysis, since the system performance reliability was based on microbial activity. However, this is more problematic for the analysis of the denitrifying biofilters, since the system performance metric was nitrate removal rather

than microbial activity. Therefore, system performance failures can occur that are independent of changes in the microbial community. For example, a decreased hydraulic retention time in a biofilter will result in less nitrate removal although the microbial activity may remain unchanged. In this case, a failure would be observed if nitrate removal is used as the performance metric but not if microbial activity is used instead. Additionally, this failure may not coincide with a change in microbial community composition. This highlights the challenge of modeling the system performance reliability based solely on microbial population dynamics if environmental and operational conditions are causing the system failure.

I believe that the ecosystem reliability model can be improved by studying the reliability of the microbial populations in isolation from the varying environmental conditions. One way to isolate the microbial community from these influences is to obtain microbial activity information from batch assays. This activity information can be used to quantify failures of the capability of the microbial community to perform its function. Since the assay environment is controlled and consistent, any changes in activity can be attributed to changes in microbial community structure. Therefore, the reliability of microbial activity can modeled using only microbial community dynamics. This approach was used for the external dataset analysis in this thesis, in which the performance metric was phenol degradation rate from batch assays.

While batch assays give valuable information on microbial activity potential, ultimately the goal of modeling an engineered ecosystem is to predict its reliability under typical operating conditions. Therefore, I believe that further exploration into approaches that incorporate changing environmental or operational conditions that directly affect

system performance into the reliability model, such as predicting the probability of failure directly caused by extreme conditions, will greatly enhance its usefulness in analyzing and predicting engineered ecosystem stability.

#### 5.1.2.2 Use of Relative Abundance Data

Another challenge unique to the analysis of engineered ecosystems is the common use of relative abundance data in microbial ecology. Since it is common practice to account for biases in the microbial analysis methods by normalizing electropherogram peak areas, the population dynamics information becomes relative abundance rather than total abundance. While this initial analysis of the denitrifying biofilters considered only relative abundance data, it is possible to discern changes in the total abundance by using quantitative polymerase chain reaction (qPCR)(Heid et al. 1996) to scale community composition results. This would improve the reliability analysis, as the difference between using relative and total abundance data can have an effect on the analysis results. In a denitrification system, where environmental conditions are ideal for denitrifying microorganisms, these microorganisms would be selected for in the system. This would lead to an overall increase in the amount of denitrifying microorganisms in the system, and subsequently improve the denitrifying activity of the system as a whole. In this situation, the relative abundance of populations possessing the *nosZ* gene in the system may not accurately explain the improved system performance, while using total abundance of denitrifying microorganisms would. This has been seen in other engineered biological treatment systems in which a positive correlation between total functional microbe abundance and system performance was demonstrated (Crocetti et al. 2000; He et al. 2010). Analyzing relative abundance values of a functional gene considers the competition

that occurs in the system between the microorganisms possessing that specific gene, but not the competition between these microorganisms as a group and all other microorganisms in the ecosystem. I believe that future reliability analysis of engineered ecosystems should incorporate methods that allow a more comprehensive understanding of the dynamics of the microbial populations present in the ecosystem. This more complete understanding of the dynamics of the microbial populations present in the ecosystem would lead to a better reliability model.

#### 5.1.2.3 Unknown System Component Configuration

In most traditional applications of reliability engineering methods, the reliability of components in a system are determined experimentally and then configured in such a way that meets the functionality and reliability needs of the system. While this may involve determining the amount of functional redundancy required to achieve the desired system reliability, the number of required operational components is known. In our engineered ecosystem reliability model, and more generally in ecological engineering, the required number of operational components and the configuration of components in the system are both unknown.

The first step in determining the configuration of microbial components in the system is to find populations with a correlation between their abundance and the system performance. This allows the user to determine whether the presence of a given population is associated with good or poor performance, and therefore whether the presence or absence of the population is considered an operational component. The idea of 'nuisance' organisms whose absence from the system is desired is one example of a population whose reliability depends upon its absence rather than its presence. This is mirrored in ecology,

where the negative impacts of nuisance organism populations are well established in both natural systems (Baskin 1998; Mack et al. 2000; Paerl et al. 2003; Clavero and García-Berthou 2005; Kindlmann et al. 2011; Vila et al. 2011) and engineered systems (Cech and Hartman 1993; Kalyuzhnyi and Fedorovich 1998; Mamais et al. 1998; Madoni et al. 2000; Wanner et al. 2000). Since not all microbial populations will express functional genes that are in their genome, and even those that do may perform the function at different rates, it is likely that in an ecosystem with limited resources where microbial populations are in competition that some functional populations will actually be nuisances to optimal system performance. Since the definition of an operational component cannot be made for microbial populations that are not correlated with good or poor performance, these populations were omitted from my analysis.

After determining whether the presence or absence of a population is considered an operational component, a block diagram must be created from these microbial components that models the system performance reliabilty. In the thesis, the system configuration was simplified by only considering microbial populations capable of performing the system function. For the external dataset and field biofilter dataset analyzed in this thesis, this came in the form of RFLP analysis of the phenol hydroxylase gene and T-RFLP analysis of the *nosZ* gene, respectively. Although total bacteria data from ARISA or T-RFLP analysis can be used to determine populations associated with good and poor performance, how these populations fit into the system reliability block diagram is not clear. There might be peripheral support roles that populations can play in the system that may lead to improved performance without possessing the functional genes responsible for the designed system function. An example of this is the proposed role that fungal populations play in the

denitrifying biofilters breaking down woodchips into carbon compounds that are ultimately utilized by denitrifying microorganisms. If certain bacterial or fungal populations are more active in these support roles, their presence in higher abundance values might be able to be included in the reliability analysis to create a more comprehensive model. In this thesis, an initial attempt of modeling denitrifying biofilter reliability was made using only the *nosZ* functional gene microbial data. Including system components contributing to the overall system function through these support roles may lead to a more robust reliability model and should be a focus of future exploration.

#### 5.1.2.4 Microbial Population 'Life Testing'

Unlike traditional reliability analysis in which life test data is provided by identical units that are tested until failure, in our methodology all of the life test data must come from multiple failures in a single time series. This raises the concern whether these 'times to failure' are comparable, because they occurred at different times, when the environment may not be similar. However, since the goal is to analyze the reliability of the system while operating under design conditions in the field, all of the individual life tests are similar in that they fall under the design operating conditions. Therefore, these individual life tests are considered comparable and can be grouped together to determine a reliability function for each microbial population.

The number and lengths of the 'times to failure' from these life tests are also highly dependent on whether the microbial population dynamics are modeled as a linear or stepwise function. In the analysis presented in this thesis I chose to use linear interpolation between points, as I felt it more accurately describes the temporal changes in the data. One result from using linear interpolation of abundance values is that if a measured abundance

value at one of the points in the time series is slightly above or below the threshold used, a very short 'time to failure' can result. This is seen in the FP07 Group P failure data, resulting in a reliability function that is much different in shape than the System or Group G functions.

## 5.2 Results from Reliability Analysis of External Dataset

Results from the reliability analysis of the external dataset show very promising results for both the reliability methods and the EcoReliAnT software. This dataset was nearly ideal for the reliability engineering analysis, as there were clear changes in microbial population structure that coincided with changes in microbial functional performance. Also, the phenol degradation rate data from batch assays provided a direct measurement of the capabilities of the microbial community present in the system. The clear relationship between certain functional microbial populations and system performance allowed for groups associated with good and poor performance to be easily formed and for the determination of threshold values that captured the trends in microbial population dynamics. One caveat is that in RFLP analysis, each microbial population is likely to produce multiple bands on the gel. In this thesis, each band was viewed as a separate population. Since the members of the microbial population groups that were formed displayed such similar population dynamics, removing one population from the group would not affect the group reliability function much. However, when populating the *k*-out-of-*n* structures with components, losing a member from a microbial population group would decrease the number of components in the system.

The external dataset analysis revealed that the reliability functions of the microbial population groups alone did not accurately model the system performance reliability. Comparing the SSE values for the Case III model, which includes both G and P to the values for a 3-out-of-3 Group G structure, we can see that including the Group P structure greatly improves the reliability model (SSE of 0.07 with and 34.86 without). This suggests that accounting for both populations associated with good and poor system performance results in a more accurate reliability model. This structure reflects the competition between microorganisms for resources and space.

One challenge with analyzing results from completely new approach, such as the ecosystem reliability analysis approach in this thesis, is that there are no other results for comparison. The interpretation and significance of these results are uncertain. Although the Case III configuration most accurately modeled the system performance reliability, the reliability function for all 5 cases were more similar to the system performance reliability than the Group G and P functions were alone. Cases I and II also demonstrated that a relatively accurate model could be created using a combination based on a single reliability function. Conceptually, if three functions are dissimilar but are all of the same form (Weibull in this case) and magnitude (all from the same time series), it is likely that some combination of one or two of the functions can be found that will be similar to the third. This raises the concern that the best-fit models were coincidental. To fully address this concern, I believe the analysis approach used in this thesis will have to be applied to several more datasets, allowing for comparison. While the resulting best-fit models for analyses of other datasets will certainly differ, I believe it would provide insight into the

validity of the approach, in particular the component configurations and the use of 'nuisance' populations.

Regardless of the interpretation of the results, the application of the ecosystem reliability analysis methods and software tools to this external dataset acted as a 'proof of concept', demonstrating the capabilities of these techniques using a real dataset. The EcoReliAnT software was used to determine reliability functions for microbial populations and model engineered ecosystem performance as a configuration of functional microbial components. Without the development of EcoReliAnT, the determination of the best-fit reliability models would be nearly impossible.

## 5.3 Results from Reliability Analysis of Field Dataset

I also applied our reliability analysis methods to more complex microbial community data from a field denitrifying biofilter. Although we propose the denitrification system can be modeled as a multi-step system (Figure 9), the current analysis considers only the *nosZ* gene, using one step to model the reliability of the entire system. Two unique challenges arose when analyzing the field biofilter data: the potential outside influences on the system performance and the use of relative abundance values for microbial populations. The implications of these challenges on the analysis have been previously discussed.

When microbial populations possessing the *nosZ* gene in the FP07 biofilter were grouped based their correlation to performance, one of the group reliability functions alone matched system performance reliability fairly well (Figure 27). However, I wanted to replicate my procedure from the analysis of the external dataset to determine if a more

accurate model could be found. Case V, consisting of a 4-out-of-8 system with 2 Group G components and 6 Group P components, resulted in a model reliability function most similar to the system reliability function. This case assumes some level of substitutability between the presence of good microbial populations and the absence of nuisance populations. Conceptually, this assumption is rooted in the idea of a competitive system in which a failure of one population can be a result of the success of another. This idea is consistent with the use of relative abundance, where one population must decrease in order for another to increase, and may explain why Case V resulted in the best reliability model.

Like the external dataset, the results from this analysis suggest that accounting for both populations associated with good and poor system performance results in a more accurate reliability model. This analysis once again demonstrated the capability of these newly developed methods to model a full-scale system. The experience gained in this analysis along with the external dataset analysis provide valuable insights into the capabilities of these methods along with the challenges associated with this novel analysis approach.

## 5.4 Results from Analysis of Laboratory Biofilters

The laboratory denitrifying biofilters that were designed, built, and operated in this thesis were successful at removing nitrate from synthetic tile drainage. The highest nitrate removal observed in the laboratory scale denitrifying biofilters occurred in R1 during operation at a low water level, when approximately 55% nitrate was removed in 8.84 hours (including up to 3 hours in the overflow and drain section). While this level of

removal is a bit higher than was seen in a study of a field biofilter by Chun and colleagues (2010), their 47% nitrate removal occurred with a hydraulic residence time of only 4.4 hours. The small discrepancy between the two is likely caused by a combination of differences in influent composition, available carbon, and dissolved oxygen concentration between the laboratory and field biofilters. Overall, the lab biofilters provide a valuable platform upon which future studies on denitrifying biofilter stability can be conducted.

In this thesis, the laboratory denitrifiying biofilters provided three particularly valuable insights. The first of these was that the biofilter performance was stable under constant conditions, but changes in environment had a major effect on system performance. The second was that denitrifying enzyme assays revealed an increased denitrifying potential of woodchips deeper in the biofilter while demonstrating a discrepancy between nitrate removal rates and denitrification rates. The third was that the biofilters exhibited an immediate change in performance in response to a change in operation. In this section I will discuss each of these points in detail.

### 5.4.1 Biofilter Performance Stability in Constant and Changing Environments

The laboratory biofilters demonstrated that under constant operating conditions woodchip biofilters can exhibit very stable performance (Figure 33). This suggests that the changes in performance observed in the field biofilters are most likely associated with changes in environmental or operational conditions. By varying things such as hydraulic retention time, temperature, or available carbon substrates, biofilter performance can be affected in two ways: directly affecting denitrification rates and indirectly affecting denitrification by causing changes in the denitrifying microbial community. The direct impact of changing environmental conditions on denitrification was highlighted by three

separate events during the period of stable performance. On 7/7/11 the tube feeding the R1 biofilter became clogged, lowering influent flow rate. This coincided with an increased nitrate removal efficiency for that day. When the flowrate was corrected, the performance returned to its previous level. On 8/30/11, woodchips were added to the top of the biofilter to replace those that were taken for samples. This caused a wood 'dust' to fall into the biofilter water and a water sample taken later that day revealed a much higher nitrate removal rate, suggesting that the system is carbon limited. Additional evidence for the influence of carbon content on performance was seen during the startup period when the carbon content of the synthetic tile drainage was forty times higher than the amount thereafter. During this period, nearly 100% nitrate removal was observed (Figure 31). Between 10/17/11 and 10/20/11 the normal laboratory room temperature of 20 degrees Celsius was elevated to 27 degrees Celsius. In this period, both biofilters exhibited approximately 20% increased nitrate removal efficiency (Figure 34). Performance again returned to its original level when the original room temperature was restored. This suggests that operating temperature can have a large effect on the denitrification rate in the biofilters.

All of these influences show a direct effect of environmental conditions on biofilter performance, and stress the importance of further exploration into methods of incorporating these influences into the biofilter reliability model. While the microbial community analysis for the laboratory biofilters has not yet been completed, the performance dynamics seen during these events were too rapid to be caused by a change in microbial community structure. As discussed earlier, a reliability model based solely on

microbial population dynamics is therefore unlikely to accurately predict these types of rapid changes in performance.

### 5.4.2 Denitrifying Enzyme Assay Results

Denitrifying enzyme assays provided information on the denitrifying potential of the microbes in the biofilters. In the modified assay method that I used in this thesis, I collected data on both nitrate removal, which is relevant in understanding biofilter system function, and nitrous oxide production, which reveals the amount of nitrate removed through denitrification. Therefore, if denitrification is the main mechanism for nitrate removal in the biofilters, these rates should be similar. However, the results from the denitrification enzyme assays showed an order of magnitude discrepancy between the amount of nitrate removed (0.0069-0.01546 mg N (dry g woodchip)$^{-1}$ h$^{-1}$) and the amount of nitrous oxide produced (0.00063-0.00176 mg N (dry g woodchip)$^{-1}$ h$^{-1}$). This suggests that either the nitrous oxide production rates do not accurately reflect the denitrification that is occurring in the assays or nitrate is being removed via some other mechanism.

Since no nitrite was detected when analyzing the water samples using ion chromatography, incomplete denitrification is unlikely to be causing this discrepancy. While some of the nitrous oxide would be dissolved in the liquid, this would also not account for an order of magnitude discrepancy. Acetylene may have not completely blocked the reduction of nitrous oxide to dinitrogen gas, resulting in lower nitrous oxide production measures. However, the nitrous oxide production rates of 0.00063-0.00176 mg N (dry g woodchip)$^{-1}$ h$^{-1}$ demonstrated in this thesis are similar to those reported previously (0.0003416-0.001433 mg N (dry g woodchip)$^{-1}$ h$^{-1}$) (Moorman et al. 2010),

suggesting that either the assay is flawed or my nitrous oxide production results accurately reflect the denitrification in the assays.

It is also possible that nitrate is being removed via another mechanism. While nitrate can be utilized for cell growth, I would not expect this to explain the amount of nitrate removal seen in these assays. If ammonium was present in the assays, nitrate may have been reduced to nitrite, which was subsequently converted to dinitrogen gas along with the ammonium via the anammox process. While it is possible that organic nitrogen in the woodchips was converted to ammonium, I do not believe this would have occurred at a rate that would affect the assay results. Nitrate may also have adsorbed to the woodchips or glass bottle, resulting in the high nitrate removal rates in the assays. However, because the assay method used in this thesis was modified to include nitrate removal data, removal rates in the assays and laboratory biofilters can be compared directly. This shows that nitrate was removed at a much higher rate in the assays (0.0069-0.01546 mg N (dry g woodchip)$^{-1}$ h$^{-1}$) than in the laboratory biofilters (0.000430-0.003006 mg N (dry g woodchip)$^{-1}$ h$^{-1}$ ). If nitrate was being removed via a mechanism other than denitrification in the assays, I would also expect this to be seen in the laboratory biofilters, but it is not. Therefore, I believe that the discrepancies between nitrate removal rates and nitrous oxide production rate needs to be explored further before any conclusions can be drawn comparing the assays and the laboratory biofilters.

While the cause of the discrepancy between the nitrate removal and nitrous oxide production rates is uncertain, the trends are similar in both rates. Of the 10 assays performed, woodchips from the deep ports in R1 removed more nitrate than those from its shallow port 7 times, while this occurred in 9 assays for R2. Nitrous oxide measurements in

these assays also showed higher nitrous oxide production rates in the deep ports in every assay analyzed. This suggests that the microorganisms deeper in the woodchip bed have a higher denitrification capability.

However, differences in moisture content between the two sampling depths would affect both of these rates, as they are normalized per dry gram of woodchip in the sample. This effect is even greater during the modified operation as the woodchips from the R1 shallow port dried out. Taking samples and replacing woodchips caused the entire R1 shallow sampling port to be filled with woodchips that were never submerged after 4 weeks of operation at low water level. This can be seen in the moisture content data (Figure 36), where there is a sudden drop in the R1 shallow port measurements. It is also interesting that the 3 weeks in which the R1 shallow port had a higher nitrate removal than its deep port were shortly after the drop in water level. I hypothesize that woodchip degradation caused the increase in the R1 shallow port nitrate removal rate in the 4 weeks following the lowering of the water level. Since no additional carbon source was used in the denitrifying enzyme assays, all of the carbon must come from woodchip decomposition. I believe that there was more available carbon on these newly exposed woodchips than in the submerged woodchips, which led to more nitrate removal in the assays. This is supported by a field biofilter study that demonstrated 75% woodchip loss in areas that were intermittently submerged and only 20% loss in areas always submerged (Moorman et al. 2010). Woli and colleagues also noticed that there was more nitrate removal in field denitrifying biofilters in the first high flows after dry periods, hypothesizing that there was more available carbon (2010). In our denitrifying enzyme assays, when the woodchips in the sample ports that were previously submerged were depleted after 4 weeks and slowly

replaced with woodchips that had never been submerged, the nitrate removal in the shallow ports dropped and remained below that seen in the deep ports.

### 5.4.3 Response to Change in Operation

While the motivation behind the change in operation of one laboratory biofilters was to achieve different levels of microbial diversity, the response to this change was unexpected. Lowering the water level on the R1 biofilter resulted in an immediate increase in nitrate removal, although the theoretical hydraulic retention time remained the same. As with the direct influences on nitrate removal discussed earlier, this operational change also highlights a change in performance that was too rapid to be caused by a change in microbial community structure. Because of this unexpected response, my focus shifted from achieving different levels of diversity to trying to explain the cause of this change in performance.

I hypothesize that the change in performance was caused by a combination of factors, including differences in hydraulic retention times and dissolved oxygen gradients between the two reactors, and a difference in denitrification activity between the shallow and deep sections of the woodchip bed. Although the theoretical hydraulic retention times in the two reactors are identical, the tracer study results indicate that the average hydraulic residence time in R2 was less than R1 (7.55 hours versus 8.84 hours)(Figure 37), which would lead to more denitrification in R1. However, the 15% difference in hydraulic residence time alone does not account for the 50% difference in nitrate removal efficiency between the two reactors. The lower water level may also have created different dissolved oxygen gradients in the woodchip beds that affected the system performance both directly and indirectly. Since denitrifying microorganisms prefer oxygen over nitrate as an electron

acceptor, higher dissolved oxygen levels in the R2 biofilter water that is closer to atmospheric conditions would result in less nitrate reduction. Additionally, I hypothesize that when both reactors were operated with high water levels that the bottom portion of each reactor was anoxic, creating an ideal environment that selected for denitrifying microorganisms. Therefore, when the water level was lowered in R1, the water was in contact with these woodchips whose microbial community was enriched with denitrifiers for a longer period. This hypothesis is supported by the denitrifying enzyme assay results discussed earlier that showed more denitrification potential in the deeper ports.

To further examine the cause of the difference in performance between the two biofilters operating at different water levels, I suggest regularly taking both chemical oxygen demand and dissolved oxygen measurements from each reactor. The chemical oxygen demand measurements should be taken at different locations along the length of the reactor to see if carbon is being depleted in the system, while dissolved oxygen measurements should be taken at different depths to determine the effect of lower water level on the oxygen gradient in the biofilter. Both of these will shed light on possible factors affecting nitrate removal in the biofilters.

# Chapter 6: Conclusions and Future Work

With a better understanding of the relationships between the functional or performance stability and the microbial community stability of engineered ecosystems, we will be able to more accurately predict failures of system performance and improve system operation and design.

While reliability methods are commonly used to analyze engineered systems, to my knowledge this thesis is the first attempt to do so with an engineered ecosystem. The goals of this research were to develop tools to analyze engineered ecosystem stability, to demonstrate the use of these tools on microbial population dynamics data from engineered ecosystems, and to operate laboratory reactors that provide information for future reliability analysis.

The Ecosytem Reliabilty Analysis Tool (EcoReliAnT) that was developed in MATLAB® proved to be very beneficial for applying both new and old reliability methods to engineered ecosystem data. While the methods proposed to analyze engineered ecosystems in this thesis could potentially be done without automation, the time required to do so would be a hindrance. The EcoReliAnT software provides a user-friendly interface, allowing powerful and rapid analysis of engineered ecosystem data.

Results from the reliability analysis of the external dataset demonstrated the capabilities of both the reliability methods and the EcoReliAnT software. The reliability of an engineered ecosystem's performance was accurately modeled as a configuration of functional microbial components. Accounting for microbial populations associated with good and poor performance in the reliability analysis led to a more accurate model.

Results from the reliability analysis of the FP07 field biofilter dataset demonstrated another application of the reliability analysis methods and software developed in this thesis. An accurate reliability model for the system was created as a configuration of microbial components possessing the *nosZ* gene. This more complex dataset also raised the important issue of outside influences on system performance.

Both of these reliability studies demonstrate that the methods and tools presented in this thesis can model engineered ecosystems using functional microbial population dynamics. Since this was the first application of these methods, the experience and results revealed possible avenues for improved functionality or model accuracy. Many of these potential improvements have already been discussed. In my opinion, the most important issue is to incorporate the direct effects of environmental and operations conditions on system performance. Another potential improvement would be to model the engineered ecosystem as a renewal process, analyzing system availability rather than reliability. Currently the reliability analysis is not dynamic, meaning that when microbial populations and engineered ecosystems fail, their function is never restored. This is unlike the actual populations and ecosystems that are being modeled, which may fail but can subsequently recover. The 'time to recovery' after a failure may actually be just as useful in engineered ecosystem design and operation as the 'time to failure'. While the Ecosystem Reliability Analysis Tool is not currently equipped for this type of analysis, much of the functionality that is in place could be used to further develop the software to incorporate renewal processes.

Results from the operation of laboratory denitrifying biofilters showed that under constant operating conditions, denitrifying biofilters exhibit very stable performance,

suggesting that periods of instability observed in the field biofilters may be caused by changing environmental or operational conditions. After lowering the water level in one of the biofilters, an immediate increase in nitrate removal was shown. Although a complete explanation for this change in performance is not yet known, this demonstrates the effect that changing operational and environmental conditions can have on system performance while highlighting the need for a greater understanding of the factors affecting denitrifying biofilter performance.

In addition to the experiments described earlier to explain the discrepancy in the denitrifying enzyme assay results and the unexpected improvement in biofilter performance that occurred when the water level was lowered, I recommend several laboratory biofiler experiments to further explore ecosystem stability concepts. The goal of the first experiment is to provide a varied environment in one reactor in an attempt to achieve differing levels of diversity. This experiment, performed by varying water level in one biofilter to provide intermittent submergence of a section of the woodchip bed, is currently underway. Once different levels of diversity are achieved, studying the stability of the system will provide a greater understanding of the relationship between functional redundancy and performance stability. Microbial community analysis during this experiment will also test the hypothesis proposed by Matthew Porter (2011), that moisture content drives organic carbon gradients and subsequently microbial community structure.

The second experiment I propose would explore the effects of environmental and operational conditions on the biofilter microbial community and system performance. This can be done by changing the operational or environmental conditions to mimic those experienced by a biofilter in the field. Operational and environmental parameters of

interest include influent flow rate, hydraulic residence time, influent nitrate concentration, and temperature. To accurately assess the effect of these changes on system performance, I suggest studying each of them individually. The results from these experiments would be valuable for several aspects of the project. It would enhance the understanding of the seasonal and spatial differences in microbial community structure that were demonstrated in the field biofilters by Matthew Porter (2011). The response of the laboratory biofilters to changes in operational or environmental conditions can also be used to improve the denitrifying biofilter reliability model. Having a better understanding of how these changing conditions directly affect the system performance will allow factors other than microbial community structure to be incorporated into the reliability analysis, leading to a more accurate reliability model. Additionally, how individual microbial populations respond to changing environmental or operational conditions and the consequences these changes in microbial community structure have on denitrifying activity and system performance may present avenues of optimizing system design or performance.

The results discussed in this thesis have contributed to the progress of this new application of reliability techniques. The Ecosystem Reliability Analysis Tool that I have developed provides a user-friendly interface for ecosystem reliability analysis while establishing a foundation for future software development. Additionally, I expect that the laboratory biofilters will shed light on the relationships between and factors that affect microbial community stability, system functional stability, and microbial functional diversity. While my research has been split between the reliability analysis tools and the laboratory biofilters, these two areas are converging to improve our understanding of engineered ecosystem stability. Having a better understanding of the factors that affect the

functional stability of denitrifying biofilters and other engineered ecosystems will ultimately allow for improved system design and operation and enhance the environmental services that these systems provide.

# References


Adamides, E., Stamboulis, Y., and Varelis, A. (2004). "Model-based assessment of military aircraft engine maintenance systems." *J. Oper. Res. Soc.,* 55(9), 957-967.

Anderson, T. W., and Darling, D. A. (1952). "Asymptotic theory of certain “goodness-of-fit” criteria based on stochastic processes." *Ann Math Stat,* 23 193-212.

Andrus, J. M. (2010). "Development and Application of Microbial Community Analysis Techniques for a Denitrifying Biofilter." Ph.D. dissertation, University of Illinois at Urbana-Champaign, Urbana, IL.

Ayala-del-Rio, H., Callister, S., Criddle, C., and Tiedje, J. (2004). "Correspondence between community structure and function during succession in phenol- and phenol-plus-trichloroethene-fed sequencing batch reactors." *Appl. Environ. Microbiol.,* 70(8), 4950-4960.

Barlow, R., and Heidtmann, K. (1984). "Computing K-Out-Of-N System Reliability." *IEEE Trans .Reliab.,* 33(4), 322-323.

Baskin, Y. (1998). "Winners and losers in a changing world: Global changes may promote invasions and alter the fate of invasive species." *Bioscience,* 48(10), 788-792.

Belfore, L. (1995). "An O(n-Center-Dot(log(2)(n))(2)) Algorithm for Computing the Reliability of K-Out-Of-N-G and K-To-L-Out-Of-N-G Systems." *IEEE Trans. Reliab.,* 44(1), 132-136.

Blattel, C. R., Williard, K. W. J., Baer, S. G., Schoonover, J. E., and Zaczek, J. I. (2009). "Ground Water Nitrogen Dynamics in Giant Cane and Forest Riparian Buffers." *Castanea,* 74(3), 259-270.

Blowes, D. W., Robertson, W. D., Ptacek, C. J., and Merkley, C. (1994). "Removal of agricultural nitrate from tile-drainage effluent water using in-line bioreactors." *J. Contam. Hydrol.,* 15(3), 207-221.

Braker, G., Zhou, J., Wu, L., Devol, A., and Tiedje, J. (2000). "Nitrite reductase genes (nirK and nirS) as functional markers to investigate diversity of denitrifying bacteria in Pacific northwest marine sediment communities." *Appl. Environ. Microbiol.,* 66(5), 2096-2104.

Braker, G., and Tiedje, J. (2003). "Nitric oxide reductase (norB) genes from pure cultures and environmental samples." *Appl. Environ. Microbiol.,* 69(6), 3476-3483.

Cameron, S. G., and Schipper, L. A. (2010). "Nitrate removal and hydraulic performance of organic carbon for use in denitrification beds." *Ecol. Eng.,* 36(11), 1588-1595.

Cech, J., and Hartman, P. (1993). "Competition between Polyphosphate and Polysaccharide Accumulating Bacteria in Enhanced Biological Phosphate Removal Systems." *Water Res.,* 27(7), 1219-1225.

Chakravarti, I. M., Laha, R. G., and Roy, J. (1967). *Handbook of Methods of Applied Statistics, Volume I.* John Wiley & Sons, New York.

Christianson, L., Castello, A., Christianson, R., Helmers, M., and Bhandari, A. (2010a). "Hydraulic Property Determination of Denitrifying Bioreactor Fill Media." *Appl. Eng. Agric.,* 26(5), 849-854.

Christianson, L., Helmers, M., and Bhandari, A. (2010b). "Bioreactor design geometry effects on nitrate removal." *ASABE - 9th International Drainage Symposium 2010, Held Jointly with CIGR and CSBE/SCGAB,* 38-44.

Christianson, L. E., Bhandari, A., and Helmers, M. J. (2011). "Pilot-Scale Evaluation of Denitrification Drainage Bioreactors: Reactor Geometry and Performance." *J. Environ. Eng.-ASCE,* 137(4), 213-220.

Chun, J. A., Cooke, R. A., Eheart, J. W., and Kang, M. S. (2009). "Estimation of flow and transport parameters for woodchip-based bioreactors: I. laboratory-scale bioreactor." *Biosyst. Eng.,* 104(3), 384-395.

Chun, J. A., Cooke, R. A., Eheart, J. W., and Cho, J. (2010). "Estimation of flow and transport parameters for woodchip-based bioreactors: II. field-scale bioreactor." *Biosyst. Eng.,* 105(1), 95-102.

Clavero, M., and García-Berthou, E. (2005). "Invasive species are a leading cause of animal extinctions." *Trends in Ecology and Evolution,* 20(3), 110.

Cooke, R. A., Doheny, A. M., and Hirschi, M. C. (2001). "Bio-reactors for edge-of-field treatment of tile outflow." ASABE paper #0120181-17.

Crocetti, G., Hugenholtz, P., Bond, P., Schuler, A., Keller, J., Jenkins, D., and Blackall, L. (2000). "Identification of polyphosphate-accumulating organisms and design of 16S rRNA-directed probes for their detection and quantitation." *Appl. Environ. Microbiol.,* 66(3), 1175-1182.

Cytryn, E., Minz, D., Gelfand, I., Neori, A., Gieseke, A., De Beer, D., and Van Rijn, J. (2005). "Sulfide-oxidizing activity and bacterial community structure in a fluidized bed reactor from a zero-discharge mariculture system." *Environ. Sci. Technol.,* 39(6), 1802-1810.

Daims, H., Purkhold, U., Bjerrum, L., Arnold, E., Wilderer, P., and Wagner, M. (2001). "Nitrification in sequencing biofilm batch reactors: Lessons from molecular approaches." *Water Sci.Technol.,* 43(3), 9-18.

David, M. B., Drinkwater, L. E., and Mclsaac, G. F. (2010). "Sources of Nitrate Yields in the Mississippi River Basin." *J. Environ. Qual.,* 39(5), 1657-1667.

Diaz, R. J., and Rosenberg, R. (2008). "Spreading dead zones and consequences for marine ecosystems." *Science,* 321(5891), 926-929.

Doheny, A. (2003). "Amelioration of tile nitrate and atrazine using inline biofilters.". M.S. thesis, University of Illinois at Urbana-Champaign, Urbana, IL.

Enevoldsen, I., and Sorensen, J. (1994). "Reliability-Based Optimization in Structural-Engineering." *Struct. Saf.,* 15(3), 169-196.

Fernandez, A., Huang, S., Seston, S., Xing, J., Hickey, R., Criddle, C., and Tiedje, J. (1999). "How stable is stable? Function versus community composition." *Appl. Environ. Microbiol.,* 65(8), 3697-3704.

Fisher, M., and Triplett, E. (1999). "Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities." *Appl. Environ. Microbiol.,* 65(10), 4630-4636.

Gavrilov, L., and Gavrilova, N. (2001). "The reliability theory of aging and longevity." *J. Theor. Biol.,* 213(4), 527-545.

Gentile, M., Yan, T., Tiquia, S. M., Fields, M. W., Nyman, J., Zhou, J., and Criddle, C. S. (2006). "Stability in a denitrifying fluidized bed reactor." *Microb. Ecol.,* 52(2), 311-321.

Gibert, O., Pomierny, S., Rowe, I., and Kalin, R. M. (2008). "Selection of organic substrates as potential reactive materials for use in a denitrification permeable reactive barrier (PRB) " *Bioresour. Technol.,* 99(16), 7587-7596.

Girvan, M., Campbell, C., Killham, K., Prosser, J., and Glover, L. (2005). "Bacterial diversity promotes community stability and functional resilience after perturbation" *Environ. Microbiol.,* 7(3), 301-313.

Goolsby, D. A., Battaglin, W. A., Lawrence, G. B., Artz, R. S., Aulenbach, B. T., Hooper, R. P., Keeney, D. R., and Stensland, G. J. (1999). *Flux and Sources of Nutrients in the Mississippi-Atchafalaya River Basin: Topic 3 Report for the Integrated Assessment on Hypoxia in the Gulf of Mexico.* NOAA Coastal Ocean Program, Silver Spring, MD.

Goseva-Popstojanova, K., and Trivedi, K. (2000). "Failure correlation in software reliability models." *IEEE Trans. Reliab.,* 49(1), 37-48.

Green, A., and Hedinger, R. (1978). "Models Relating Ultraviolet-Light and Non-Melanoma Skin Cancer Incidence." *Photochem .Photobiol.,* 28(2), 283-291.

Greenan, C., Moorman, T., Kaspar, T., Parkin, T., and Jaynes, D. (2006). "Comparing carbon substrates for denitrification of subsurface drainage water." *J. Environ. Qual.,* 35(3), 824-829.

Greenan, C. M., Moorman, T. B., Parkin, T. B., Kaspar, T. C., and Jaynes, D. B. (2009). "Denitrification in Wood Chip Bioreactors at Different Water Flows." *J. Environ. Qual.,* 38(4), 1664-1671.

Griffiths, B. S., Kuan, H. L., Ritz, K., Glover, L. A., McCaig, A. E., and Fenwick, C. (2004). "The relationship between microbial community structure and functional stability, tested experimentally in an upland pasture soil." *Microb. Ecol.,* 47(1), 104-113.

Grimm, V., Schmidt, E., and Wissel, C. (1992). "On the Application of Stability Concepts in Ecology." *Ecol. Model.,* 63(1-4), 143-161.

He, S., Bishop, F. I., and McMahon, K. D. (2010). "Bacterial Community and "Candidatus Accumulibacter" Population Dynamics in Laboratory-Scale Enhanced Biological Phosphorus Removal Reactors." *Appl. Environ. Microbiol.,* 76(16), 5479-5487.

Heid, C., Stevens, J., Livak, K., and Williams, P. (1996). "Real time quantitative PCR." *Genome Res.,* 6(10), 986-994.

Howarth, R., Billen, G., Swaney, D., Townsend, A., Jaworski, N., Lajtha, K., Downing, J., Elmgren, R., Caraco, N., Jordan, T., Berendse, F., Freney, J., Kudeyarov, V., Murdoch, P., and Zhu, Z. (1996). "Regional nitrogen budgets and riverine N&P fluxes for the drainages to the North Atlantic Ocean: Natural and human influences" *Biogeochemistry,* 35(1), 75-139.

Ines, M., Soares, M., and Abeliovich, A. (1998). "Wheat straw as substrate for water denitrification." *Water Res.,* 32(12), 3790-3794.

Jaynes, D. B., and James, D. E. (2007). "The extent of farm drainage in the United States." *ASA-CSSA-SSSA Annual Meeting Abstracts*, Nov. 4-8, 2007, New Orleans, LA.

Jaynes, D. B., Kaspar, T. C., Moorman, T. B., and Parkin, T. B. (2008). "In situ bioreactors and deep drain-pipe installation to reduce nitrate losses in artificially drained fields." *J. Environ. Qual.,* 37(2), 429-436.

Jørgensen, S. E., and Mejer, H. (1977). "Ecological buffer capacity." *Ecol. Model.,* 3(1), 39-45,47,49,51,53-61.

Kalita, P., Algoazany, A., Mitchell, J., Cooke, R., and Hirschi, M. (2006). "Subsurface water quality from a flat tile-drained watershed in Illinois, USA." *Agric. Ecosyst. Environ.,* 115(1-4), 183-193.

Kalyuzhnyi, S., and Fedorovich, V. (1998). "Mathematical modelling of competition between sulphate reduction and methanogenesis in anaerobic reactors." *Bioresour. Technol.,* 65(3), 227-242.

Kindlmann, P., Ameixa, O. M. C. C., and Dixon, A. F. G. (2011). "Ecological effects of invasive alien species on native communities, with particular emphasis on the interactions between aphids and ladybirds." *Biocontrol,* 56(4), 469-476.

Korell, J., Coulter, C. V., and Duffull, S. B. (2011). "A statistical model for red blood cell survival." *J. Theor. Biol.,* 268(1), 39-49.

Kovacic, D., David, M., Gentry, L., Starks, K., and Cooke, R. (2000). "Effectiveness of constructed wetlands in reducing nitrogen and phosphorus export from agricultural tile drainage." *J. Environ. Qual.,* 29(4), 1262-1274.

Kovacic, D. A., Twait, R. M., Wallace, M. P., and Bowling, J. M. (2006). "Use of created wetlands to improve water quality in the Midwest-Lake Bloomington case study." *Ecol. Eng.,* 28(3 SPEC. ISS.), 258-270.

Laird, R. A., and Sherratt, T. N. (2009). "The evolution of senescence through decelerating selection for system reliability." *J. Evol. Biol.,* 22(5), 974-982.

Laird, R. A., and Sherratt, T. N. (2010). "The evolution of senescence in multi-component systems." *BioSystems,* 99(2), 130-139.

Leigh, E. (1981). "The Average Lifetime of a Population in a Varying Environment." *J. Theor. Biol.,* 90(2), 213-239.

Liu, W., Marsh, T., Cheng, H., and Forney, L. (1997). "Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA." *Appl. Environ. Microbiol.,* 63(11), 4516-4522.

Mack, R., Simberloff, D., Lonsdale, W., Evans, H., Clout, M., and Bazzaz, F. (2000). "Biotic invasions: Causes, epidemiology, global consequences, and control." *Ecol. Appl.,* 10(3), 689-710.

Madoni, P., Davoli, D., and Gibin, G. (2000). "Survey of filamentous microorganisms from bulking and foaming activated-sludge plants in Italy." *Water Res.,* 34(6), 1767-1772.

Mamais, D., Andreadakis, A., Noutsopoulos, C., and Kalergis, C. (1998). "Causes of, and control strategies for, Microthrix parvicella bulking and foaming in nutrient removal activated sludge systems." *Water Sci. Technol.,* 37(4-5), 9-17.

Martienssen, M., and Schops, R. (1997). "Biological treatment of leachate from solid waste landfill sites - Alterations in the bacterial community during the denitrification process." *Water Res.,* 31(5), 1164-1170.

Mayer, P. M., Reynolds, S. K., Jr., McCutchen, M. D., and Canfield, T. J. (2007). "Meta-analysis of nitrogen removal in riparian buffers." *J. Environ. Qual.,* 36(4), 1172-1180.

Miura, Y., Hiraiwa, M. N., Ito, T., Itonaga, T., Watanabe, Y., and Okabe, S. (2007). "Bacterial community structures in MBRs treating municipal wastewater: Relationship between community stability and reactor performance." *Water Res.,* 41(3), 627-637.

Moorman, T. B., Parkin, T. B., Kaspar, T. C., and Jaynes, D. B. (2010). "Denitrification activity, wood loss, and N2O emissions over 9 years from a wood chip bioreactor." *Ecol. Eng.,* 36(11), 1567-1574.

Moranda, P. (1975). "Prediction of Software Reliability during Debugging." *Proc. Annu. Reliab. Maintainab. Symp.*, 327-332.

Moyer, C. L., Dobbs, F. C., and Karl, D. M. (1994). "Estimation of diversity and community structure through restriction fragment length polymorphism distribution analysis of bacterial 16S rRNA genes from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii." *Appl. Environ. Microbiol.,* 60(3), 871-879.

Muyzer, G., deWaal, E., and Uitterlinden, A. (1993). "Profiling of Complex Microbial-Populations by Denaturing Gradient Gel-Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes-Coding for 16s Ribosomal-Rna." *Appl. Environ. Microbiol.,* 59(3), 695-700.

Neil, M., Fenton, N., Forey, S., and Harris, R. (2001). "Using Bayesian belief networks to predict the reliability of military vehicles." *Comput. Control. Eng. J.,* 12(1), 11-20.

Nisbet, R. M., and Gurney, W. S. C. (1982). *Modelling Fluctuating Populations.* John Wiley & Sons, Chichester (UK).

O'Geen, A. T., Budd, R., Gan, J., Maynard, J. J., Parikh, S. J., and Dahlgren, R. A. (2010). "Mitigating Nonpoint Source Pollution in Agriculture with Constructed and Restored Wetlands." *Adv. Agron.,* 108 1-76.

Paerl, H. W., Dyble, J., Moisander, P. H., Noble, R. T., Piehler, M. F., Pinckney, J. L., Steppe, T. F., Twomey, L., and Valdes, L. M. (2003). "Microbial indicators of aquatic ecosystem change: current applications to eutrophication studies." *FEMS Microbiol. Ecol.,* 46(3), 233-246.

Philippot, L., Piutti, S., Martin-Laurent, F., Hallet, S., and Germon, J. (2002). "Molecular analysis of the nitrate-reducing community from unplanted and maize-planted soils." *Appl. Environ. Microbiol.,* 68(12), 6121-6128.

Pholchan, M. K., Baptista, J. d. C., Davenport, R. J., and Curtis, T. P. (2010). "Systematic study of the effect of operating variables on reactor performance and microbial diversity in laboratory-scale activated sludge reactors." *Water Res.,* 44(5), 1341-1352.

Porter, M. D. (2011). "Microbial community dynamics in denitrifying biofilters receiving agricultural drainage." M.S. thesis, University of Illinois at Urbana-Champaign, Urbana, IL.

Rackwitz, R., and Fiessler, B. (1978). "Structural Reliability Under Combined Random Load Sequences." *Comput. Struct.,* 9(5), 489-494.

Rich, J., Heichen, R., Bottomley, P., Cromack, K., and Myrold, D. (2003). "Community composition and functioning of denitrifying bacteria from adjacent meadow and forest soils." *Appl. Environ. Microbiol.,* 69(10), 5974-5982.

Robertson, W., Blowes, D., Ptacek, C., and Cherry, J. (2000). "Long-term performance of in situ reactive barriers for nitrate remediation." *Ground Water,* 38(5), 689-695.

Rosch, C., Mergel, A., and Bothe, H. (2002). "Biodiversity of denitrifying and dinitrogen-fixing bacteria in an acid forest soil." *Appl. Environ. Microbiol.,* 68(8), 3818-3829.

Rosin, P., and Rammler, E. (1933). "The Laws Governing the Fineness of Powdered Coal." *J. Inst. Fuel,* 7 29-36.

Rowan, A. K., Snape, J. R., Fearnside, D., Barer, M. R., Curtis, T. P., and Head, I. M. (2003). "Composition and diversity of ammonia-oxidising bacterial communities in wastewater treatment reactors of different design treating identical wastewater." *FEMS Microbiol. Ecol.,* 43(2), 195-206.

Ruan, Q., Dutta, D., Schwalbach, M. S., Steele, J. A., Fuhrman, J. A., and Sun, F. (2006). "Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors." *Bioinformatics,* 22(20), 2532-2538.

Rushdi, A. M. (1986). "Utilization of symmetric switching functions in the computation of k-out-of-n system reliability." *Microelectronics Reliability,* 26(5), 973-987.

Saleh, J., and Marais, K. (2006). "Highlights from the early (and pre-) history of reliability engineering." *Reliab .Eng. Syst. Saf.,* 91(2), 249-256.

Sanderson, E., and Zewde, E. (1976). "Groundwater Availability in Champaign County." *Illinois State Water Survey Circular 124*, Urbana, IL.

Schipper, L., and Vojvodic-Vukovic, M. (1998). "Nitrate removal from groundwater using a denitrification wall amended with sawdust: Field trial." *J. Environ. Qual.,* 27(3), 664-668.

Schipper, L., and Vojvodic-Vukovic, M. (2000). "Nitrate removal from groundwater and denitrification rates in a porous treatment wall amended with sawdust." *Ecol. Eng.,* 14(3), 269-278.

Schipper, L., and Vojvodic-Vukovic, M. (2001). "Five years of nitrate removal, denitrification and carbon dynamics in a denitrification wall." *Water Res.,* 35(14), 3473-3477.

Schipper, L., Barkle, G., Hadfield, J., Vojvodic-Vukovic, M., and Burgess, C. (2004). "Hydraulic constraints on the performance of a groundwater denitrification wall for nitrate removal from shallow groundwater." *J. Contam. Hydrol.,* 69(3-4), 263-279.

Schipper, L., Barkle, G., and Vojvodic-Vukovic, M. (2005). "Maximum rates of nitrate removal in a denitrification wall." *J. Environ. Qual.,* 34(4), 1270-1276.

Schipper, L. A., Cameron, S. C., and Warneke, S. (2010). "Nitrate removal from three different effluents using large-scale denitrification beds." *Ecol. Eng.,* 36(11), 1552-1557.

Smith, N., Yu, Z., and Mohn, W. (2003). "Stability of the bacterial community in a pulp mill effluent treatment system during normal operation and a system shutdown." *Water Res.,* 37(20), 4873-4884.

Snedecor, G. W., and Cochran, W. G. (1989). *Statistical Methods.* Iowa State University Press, Ames, IA.

Stamper, D., Walch, M., and Jacobs, R. (2003). "Bacterial population changes in a membrane bioreactor for graywater treatment monitored by denaturing gradient gel electrophoretic analysis of 16S rRNA gene fragments." *Appl. Environ. Microbiol.,* 69(2), 852-860.

Starr, R., and Cherry, J. (1994). "In-Situ Remediation of Contaminated Ground-Water - the Funnel-And-Gate System." *Ground Water,* 32(3), 465-476.

Stephens, M. A. (1974). "Edf Statistics for Goodness of Fit and some Comparisons." *J. Am. Stat. Assoc.,* 69(347), 730-737.

Stephens, M. A. (1976). "Asymptotic Results for Goodness-Of-Fit Statistics with Unknown Parameters." *Ann. Stat.,* 4(2), 357-369.

Stephens, M. A. (1977a). "Goodness of Fit for Extreme Value Distribution." *Biometrika,* 64(3), 583-588.

Stephens, M. A. (1977b). "Goodness of Fit Tests with Special Reference to Tests for Exponentiality." *Technical Report No. 262*, Department of Statistics, Stanford University, Stanford, CA.

Stephens, M. A. (1979). "Tests of Fit for the Logistic Distribution Based on the Empirical Distribution Function." *Biometrika,* 66(3), 591-595.

Stone, M., and Krishnappan, B. (1997). "Transport characteristics of tile-drain sediments from an agricultural watershed." *Water Air Soil Pollut.,* 99(1-4), 89-103.

Tanner, R. S. (1997). "Cultivation of bacteria and fungi." In *Manual of Environmental Microbiology,* C. J. Hurst, G. R. Knudsen, M. J. McInerney, L. D. Stetzenbach, and M. V. Walter, eds., American Society for Microbiology, Washington, D.C., 52-60.

Tees, D., Woodward, J., and Hammer, D. (2001). "Reliability theory for receptor-ligand bond dissociation." *J. Chem. Phys.,* 114(17), 7483-7496.

Tiedje, J., Simkins, S., and Groffman, P. (1989). "Perspectives on Measurement of Denitrification in the Field Including Recommended Protocols for Acetylene Based Methods." *Plant Soil,* 115(2), 261-284.

van Driel, P., Robertson, W., and Merkley, L. (2006). "Denitrification of agricultural drainage using wood-based reactors." *Trans. ASABE,* 49(2), 565-573.

Vaneechoutte, M., Debeenhouwer, H., Claeys, G., Verschraegen, G., De Rouck, A., Paepe, N., Elaichouni, A., and Portaels, F. (1993). "Identification of Mycobacterium Species by using Amplified Ribosomal Dna Restriction Analysis" *J. Clin. Microbiol.,* 31(8), 2061-2065.

Vila, M., Espinar, J. L., Hejda, M., Hulme, P. E., Jarosik, V., Maron, J. L., Pergl, J., Schaffner, U., Sun, Y., and Pysek, P. (2011). "Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems." *Ecol. Lett.,* 14(7), 702-708.

Volokita, M., Abeliovich, A., and Soares, M. (1996). "Denitrification of groundwater using cotton as energy source." *Water Sci. Technol.,* 34(1-2), 379-385.

von Canstein, H., Li, Y., Felske, A., and Wagner-Dobler, I. (2001). "Long-term stability of mercury-reducing microbial biofilm communities analyzed by 16S-23S rDNA interspacer region polymorphism." *Microb. Ecol.,* 42(4), 624-634.

Vu, K., and Stewart, M. (2000). "Structural reliability of concrete bridges including improved chloride-induced corrosion models." *Struct. Saf.,* 22(4), 313-333.

Wang, X., Wen, X., Yan, H., Ding, K., Zhao, F., and Hu, M. (2011). "Bacterial community dynamics in a functionally stable pilot-scale wastewater treatment plant." *Bioresour. Technol.,* 102(3), 2352-2357.

Wanner, J., Ruzickova, I., Krhutkova, O., and Pribyl, M. (2000). "Activated sludge population dynamics and wastewater treatment plant design and operation." *Water Science and Technology,* 41(9), 217-225.

Weibull, W. (1939). "A statistical theory of the strength of material." *Ing. Vetenskapa Acad. Handlingar,* 151 1-45.

Wildman, T. A. (2002). "Design of field-scale bioreactors for bioremediation of nitrate in tile drainage effluent." M.S. thesis, University of Illinois at Urbana-Champaign, Urbana, IL.

Wissel, C., and Stocker, S. (1991). "Extinction of Populations by Random Influences." *Theor. Popul. Biol.,* 39(3), 315-328.

Woese, C. (1987). "Bacterial Evolution." *Microbiol. Rev.,* 51(2), 221-271.

Wolda, H. (1978). "Fluctuations in Abundance of Tropical Insects." *Am. Nat.,* 112(988), 1017-1045.

Woli, K. P., David, M. B., Cooke, R. A., McIsaac, G. F., and Mitchell, C. A. (2010). "Nitrogen balance in and export from agricultural fields associated with controlled drainage systems and denitrifying bioreactors." *Ecol. Eng.,* 36(11), 1558-1566.

Woodward, K. B., Fellows, C. S., Conway, C. L., and Hunter, H. M. (2009). "Nitrate removal, denitrification and nitrous oxide production in the riparian zone of an ephemeral stream." *Soil Biol. Biochem.,* 41(4), 671-680.

Yamada, T., Logsdon, S. D., Tomer, M. D., and Burkart, M. R. (2007). "Groundwater nitrate following installation of a vegetated riparian buffer." *Sci. Total Environ.,* 385(1-3), 297-309.

Zumft, W. (1997). "Cell biology and molecular basis of denitrification." *Microbiol. Mol. Biol. Rev.,* 61(4), 533-+.

# Appendix

## Ecosystem Reliability Analysis Tool (EcoReliAnT) User's Manual

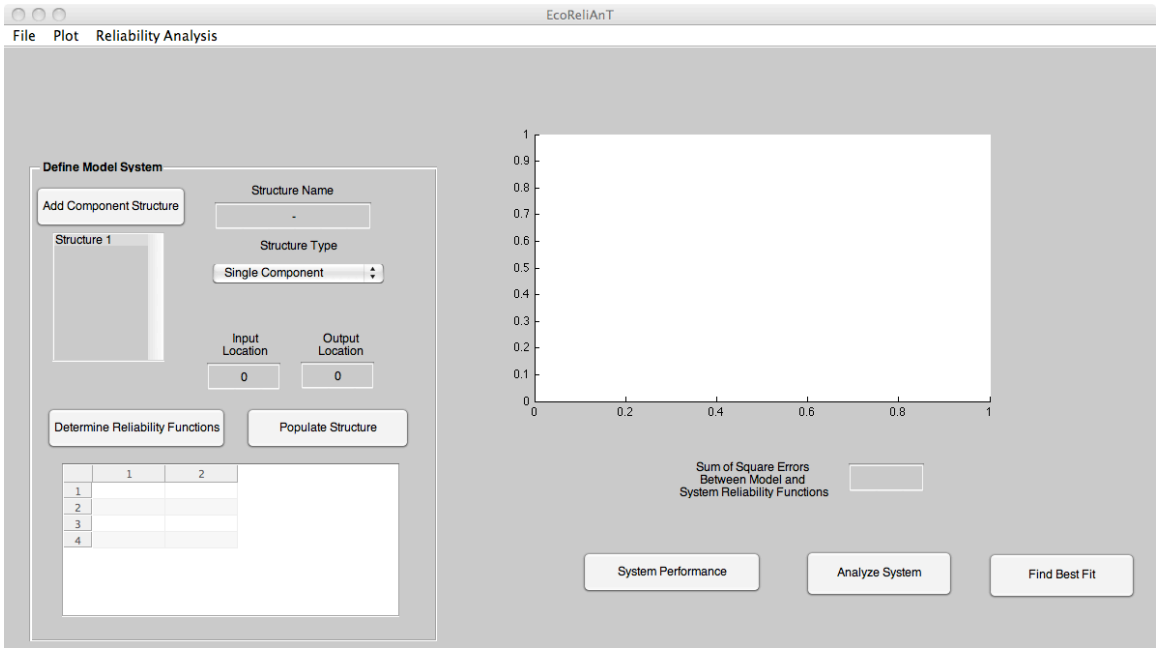



Figure 41. EcoReliAnT main window.

The Ecosystem Reliability Analysis Tool (EcoReliAnT) was created in MATLAB® to allow the user to quickly and easily analyze ecosystems using the reliability methods put forth in this thesis. When EcoReliAnT is initiated, the main window in which the user defines the system that will be analyzed is opened (Figure 41). In order to analyze a system using EcoReliAnT follow these steps:

1. Define System Configuration

    a. Add Necessary Component Structures

b. Choose Component Structure Type

c. Define Input and Output Locations of Structures to Complete System

2. Determine Reliability Functions for each of the Component Structures

a. Load data

1. Filter Data

b. Create Groups if Desired

1. Combine Input Values of Group

2. Combine Datasets of Group

c. Create Threshold Values for Determination of Presence/Absence Spans

d. Choose Options Used in Determination of Reliability Functions

1. Stepwise/Linear Interpolation

2. Use Censored Data or Not

3.Spans of Presence vs. Spans of Absence

e. Choose Reliability Model

f. When you have found fits you like, Save Fits

3. Populate Structures with Components using chosen Reliability Functions

4. Determine Reliability Function for the System based on Functional Performance

5. Analyze System Reliability

**1. Define System Configuration**

**a. Add Necessary Component Structures**

Component structures are used in EcoReliAnT to define the system configuration to be analyzed. A component structure can be made up of any components in the system that

have input and output at the same location. If the output from one component or group of components then becomes the input for another component, they need to be in separate component structures.

**b. Choose Component Structure Type**

Component Structure Types:

Single Component

Parallel Components

*k*-out-of-*n* Components

Bridge

1. In a single component structure, only one reliability function may be chosen and is used as the reliability of the entire structure.

2. In a parallel component structure, only one of the parallel components must be operational for the structure to be operational. Therefore, the reliability of this structure is the probability that at least one component is operational. There are effectively two ways to create parallel structures in EcoReliAnT. One is to use the 'parallel components' structure type option, and the other is to create completely different 'single component' structures that have identical input and output locations. The first method is easiest when the reliability functions for the parallel components can be derived from the same input

data, while the second method is easier if they are derived from different datasets. The reliability of parallel components are calculated using Equation 15.

3. In *k*-out-of-*n* systems, at least *k* components of the total *n* must be operational for the component structure to be operational. If a *k*-out-of-*n* system component structure is chosen, a box to enter the desired k value will appear. The total number reliability functions used to populate the structure determines the value of n. Therefore, if 5 functions are selected in the 'Populate Structure' interface, the total number of components (*n*) in the *k*-out-of-*n* system will be 5. The reliability of *k*-out-of-*n* systems are calculated using Equation 18.

4. Bridges are structures that are used to define more complex systems and will be discussed further later.

**c. Define Input and Output Locations of Structures to Complete System**

Where to place a node:

1. Before the first and after the last component structure
2. In between component structures in series
3. Anywhere a pathway splits between two or more component structures
4. Anywhere pathways from two or more component structures combine

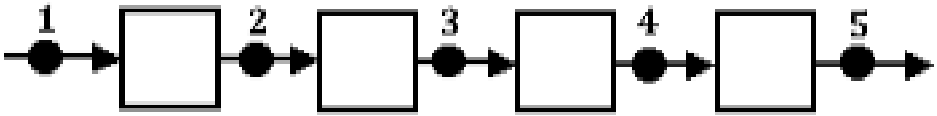


Figure 42. Simple series system with numbered nodes indicating input and output locations.

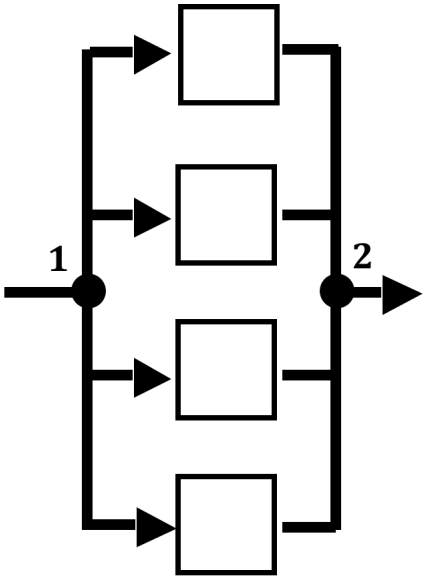


Figure 43. Simple parallel system with numbered nodes indicating input and output locations.

Bridges are used to 'bridge' the space between two nodes if there is no component structure in between. This essentially places a component that this always operational between the two nodes and is necessary for correct determination of system reliability.

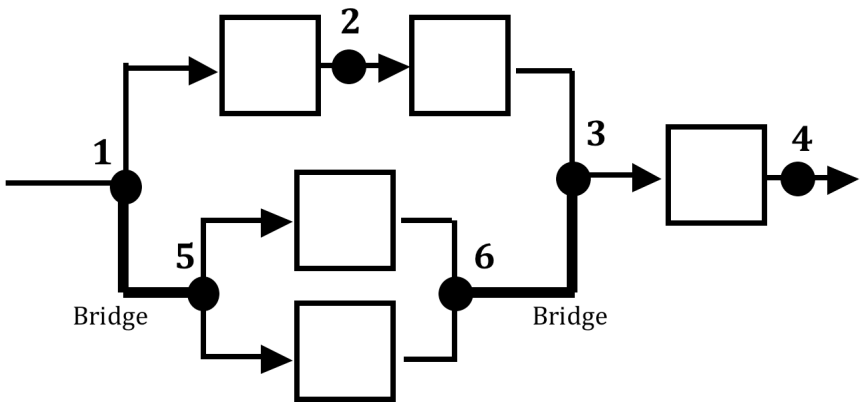


Figure 44. Complex component system with numbered nodes indicating input and output locations. Bridges are used to 'bridge the gap' between two nodes if there are no components in between them.

**2. Determine Reliability Functions for Component Structures**

To create reliability functions for microbial populations from population dynamics, highlight the structure in which the functions will be used and click the 'Determine Reliability Functions' button. This will open the Reliability Function Fit Tool (RFFT) (Figure 45).

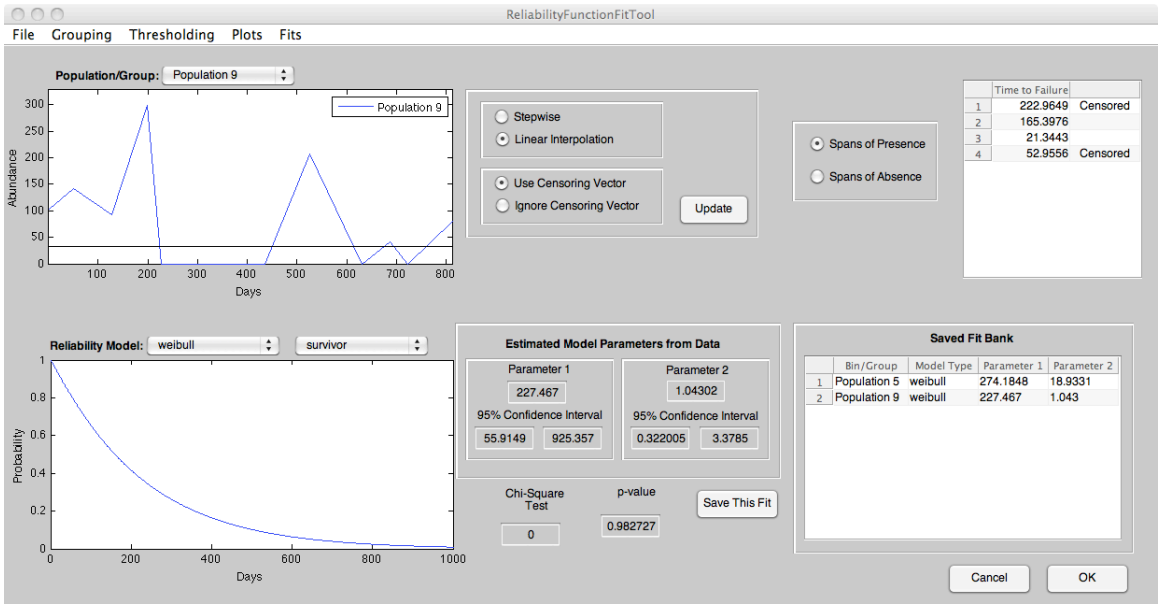
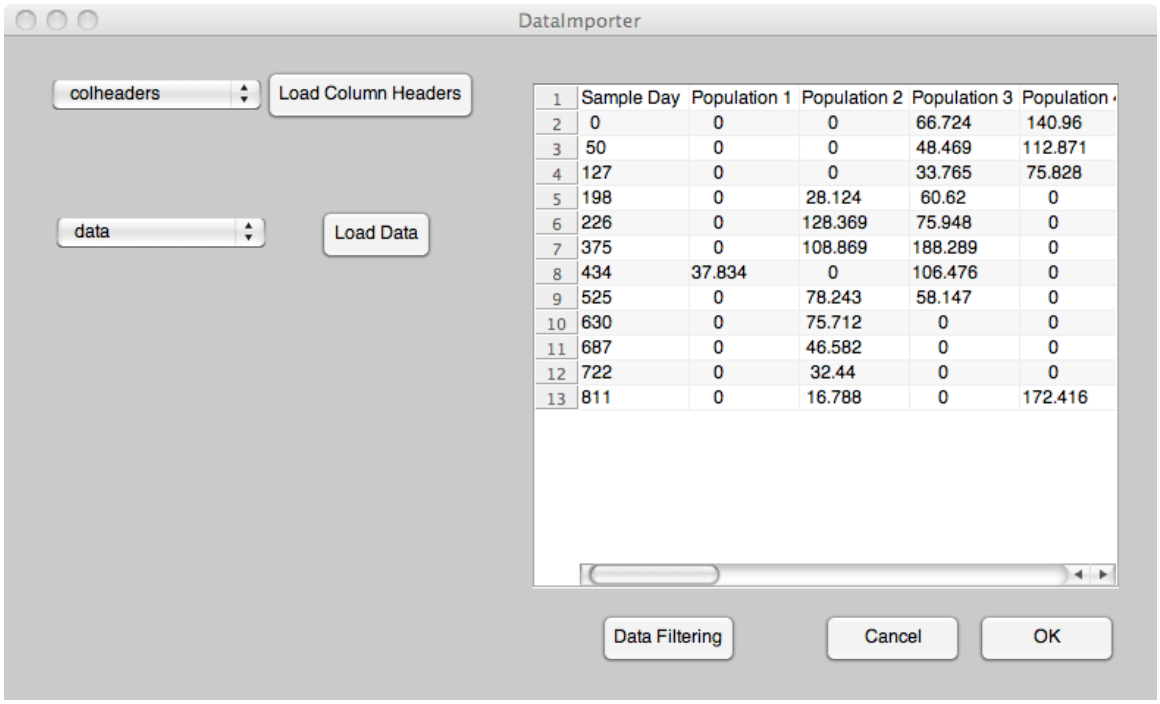



Figure 45. EcoReliAnT Reliability Function Fit Tool is used to determine reliability functions for microbial populations.

### a. Load data

To import microbial population data into the RFFT, select Import Data from the File Menu. This will open the DataImporter GUI (Figure 46). In this GUI, you can import your time series data along with the column headers from the workspace to be used throughout the program to signify the population in the corresponding column.




| | Sample Day | Population 1 | Population 2 | Population 3 | Population |
|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 66.724 | 140.96 |
| 2 | 50 | 0 | 0 | 48.469 | 112.871 |
| 3 | 127 | 0 | 0 | 33.765 | 75.828 |
| 4 | 198 | 0 | 28.124 | 60.62 | 0 |
| 5 | 226 | 0 | 128.369 | 75.948 | 0 |
| 6 | 375 | 0 | 108.869 | 188.289 | 0 |
| 7 | 434 | 37.834 | 0 | 106.476 | 0 |
| 8 | 525 | 0 | 78.243 | 58.147 | 0 |
| 9 | 630 | 0 | 75.712 | 0 | 0 |
| 10 | 687 | 0 | 46.582 | 0 | 0 |
| 11 | 722 | 0 | 32.44 | 0 | 0 |
| 12 | 811 | 0 | 16.788 | 0 | 172.416 |

Figure 46. EcoReliAnT Reliability Function Fit Tool Data Importer window which allows user to import time series data and column headers from the workspace for reliability analysis.

### 1. Filter Data

The DataImporter GUI also allows the user to filter the data before bringing it into the RFFT. This is useful for large datasets if only a portion of the populations will be used in the analysis. To filter the data before importing it, click on the 'Filter Data' button. This will open the DataFiltering GUI. In the DataFiltering GUI, data can be filtered before importing based on two criteria: a defined number of times a population must appear in the time series and a value requirement that a population must reach to be included in the imported data. Changing these values in the corresponding boxes will display the data that will be imported based on these criteria. To apply the data filter defined in the DataFiltering GUI, press 'OK'.

### b. Create Groups if Desired

One option that the RFFT give the user when determining reliability functions from microbial populations is to group populations together to create one function for the group. To create groups from your dataset, select Create Groups from the Grouping menu, opening the GroupCreator window (Figure 47).

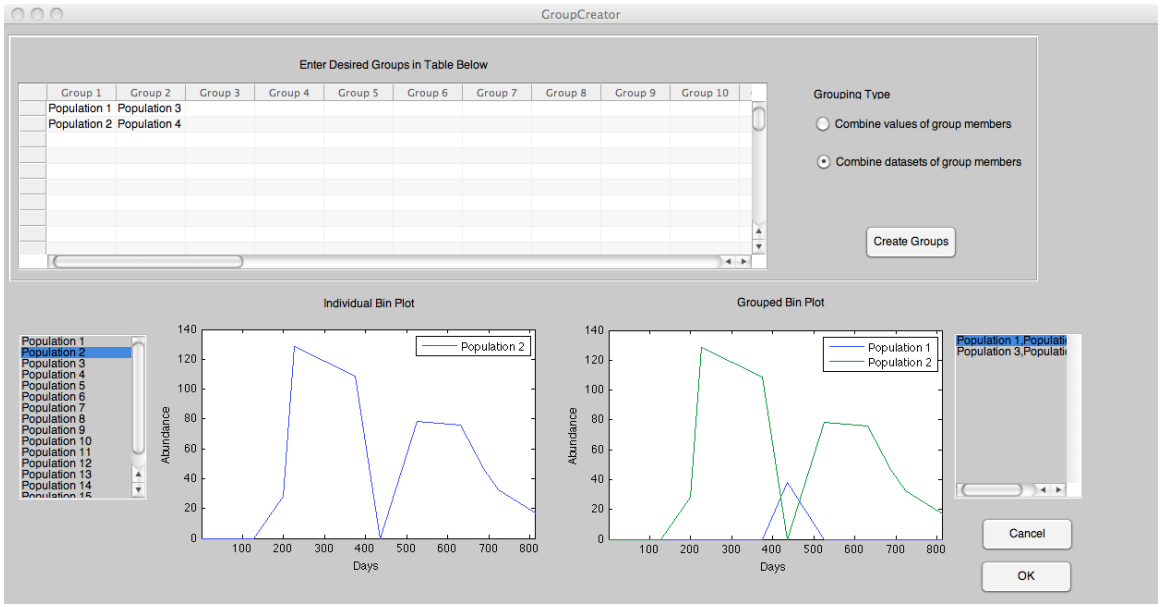



Figure 47. EcoReliAnT Reliability Function Fit Tool Group Creator window allows user to combine microbial populations into groups for future analysis.

When using groups of populations in the RFFT, there are two ways in which the data from the grouped populations can be used to determine a reliability function for the group: the time series data for the populations in the group can be combined or the population dynamics for the grouped populations can be analyzed individually and the datasets combined afterward.

If ‘Combine RA of Groups’ is selected, the data from each population in the group at each time point will be added into one time series. This time series will then be used to determine the reliability functions for the group.

If ‘Combine Group Datasets’ is selected, the population dynamics of the populations in the group are analyzed individually based on the defined thresholds, and the datasets are subsequently combined and used to create a reliability function for the group.

**c. Create Threshold Values for Determination of Spans of Absence & Presence**

The Reliability Function Fit Tool allows the user to set a threshold value for each population that determines when presence and absence occur. For example, if a relative abundance threshold of 5% is used, the spans of presence of the population will be the lengths of time the population is above 5% relative abundance. To set these thresholds, choose Create Threshold from the Thresholding menu, opening the Presence Threshold Creator window (Figure 48).

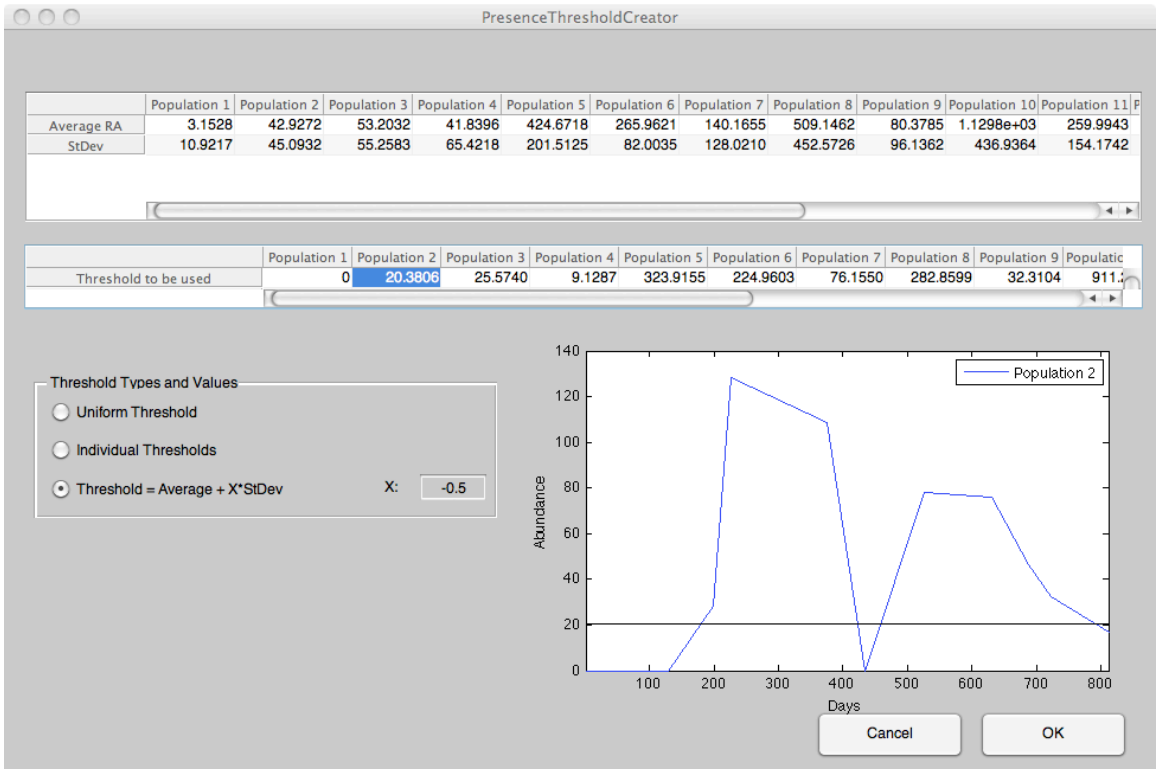



Figure 48. EcoReliAnT Reliability Function Fit Tool Presence Threshold Creator allows user to create thresholds for determination of presence absence spans for each population in the dataset.

Statistics for the microbial populations appear in the upper table. The statistics displayed are the Average Value and the Standard Deviation of each population. The bottom table displays the threshold values that are currently selected for use in the analysis. Three different types of thresholds can be chosen from the options in the window:

1. Uniform Threshold: One Threshold Value applied to all populations
2. Individual Thresholds: Allows Manual Entry of Thresholds in table above
3. Threshold= Average+X*StDev: Uses Population Statistics to Calculate a different threshold for each bin. X Value is entered into the labeled box.

Highlighting a threshold value in the bottom table will display the microbial population time series data and the currently entered threshold for that population in the chart below.

### d. Choose Options Used in Determination of Reliability Functions

The RFFT has several additional options available when determining reliability functions.

#### 1. Stepwise or Linear Interpolation

Determines whether a stepwise function is used to determine presence spans or if linear interpolation is used between data points.

#### 2. Use Censored Data or Not

If a microbial population is present in the first or last sample, the length of time until it is no longer present is known to be longer than the measure time but not by how much. This is called right hand censoring and it affects the determination of the reliability function for the dataset. The user has the option of considering the affect of the censored data or considering it as complete data. Too many censored datapoints will not allow the determination of reliability functions using the maximum likelihood estimation method.

#### 3.Spans of Presence vs. Spans of Absence

Depending on the system and microbial populations being analyzed, the user may be interested in the spans that a population is absent rather than those in which it is present. This may occur if the population is a 'nuisance' and results in failure of the system. This option allows the user to choose which span to be used in the determination of the

reliability function. Saving a Fit using 'Spans of Absence' data will automatically be noted in the name of the Fit as 'Absence of…'.

#### e. Choose Reliability Model

Selecting a reliability function from the dropdown menu will fit a function of that type to the data and display the parameters in the boxes below. A chi-square goodness-of-fit test is also run for the selected data and model type and the calculated p-value is displayed in the labeled box. This can aid in the selection of the reliability model to use in the analysis.

#### f. Save Fits

Clicking the 'Save This Fit' button will save the currently displayed reliability function to the 'Saved Fits Bank'. This and all other saved fits will then be displayed in the table at the bottom of the window.

### 3. Populate Structures with Components using Reliability Functions

Once Reliability Functions for a component structure have been defined, the selected structure must be populated. Clicking 'Populate Structure' opens a window allowing the user to select which of the previously Saved Fits from the Saved Fits Bank to use in the structure (Figure 49). Depending on the structure type, one or many fits may be used. In parallel or *k*-out-of-*n* structure types, each fit selected will be considered a parallel component. Therefore, to make multiple components with the same Saved Fit, add the fit to the select fits list multiple times.

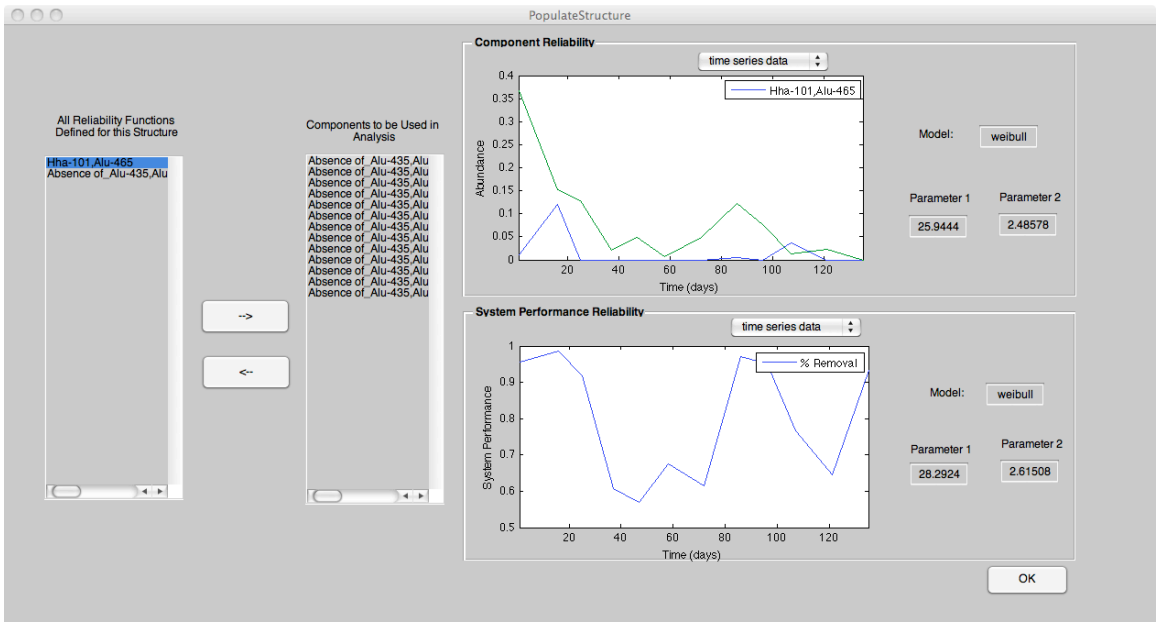



Figure 49. EcoReliAnT Reliability Function Fit Tool Populate Structure Window allows user to add or subtract components from a structure for analysis. The list box on the left shows all of the reliability functions that have been defined for the structure and the list box on the right shows the components that are current in the structure. Clicking on the left or right arrows with add or remove the highlighted reliability function. Graphs of the time series data or reliability functions for both the highlighted population and the system are displayed on the right of the window.

## 4. Determine Reliability Function for the System Performance

To open the Performance Reliability Function Fit Tool, click on 'System Performance' (Figure 50). The steps used to create a reliability function for microbial populations are repeated for the system functional performance, but they are all performed in one window with the exception of importing the data. When the desired system reliability function has been created, highlight it in the 'Saved Reliability Functions' table and click the right arrow to move it to the 'System Performance Function to be Used in Analysis' table and click 'OK'. The system reliability function can be changed at any time by

highlighting a different function and clicking the right arrow to replace the function to be used in the analysis.

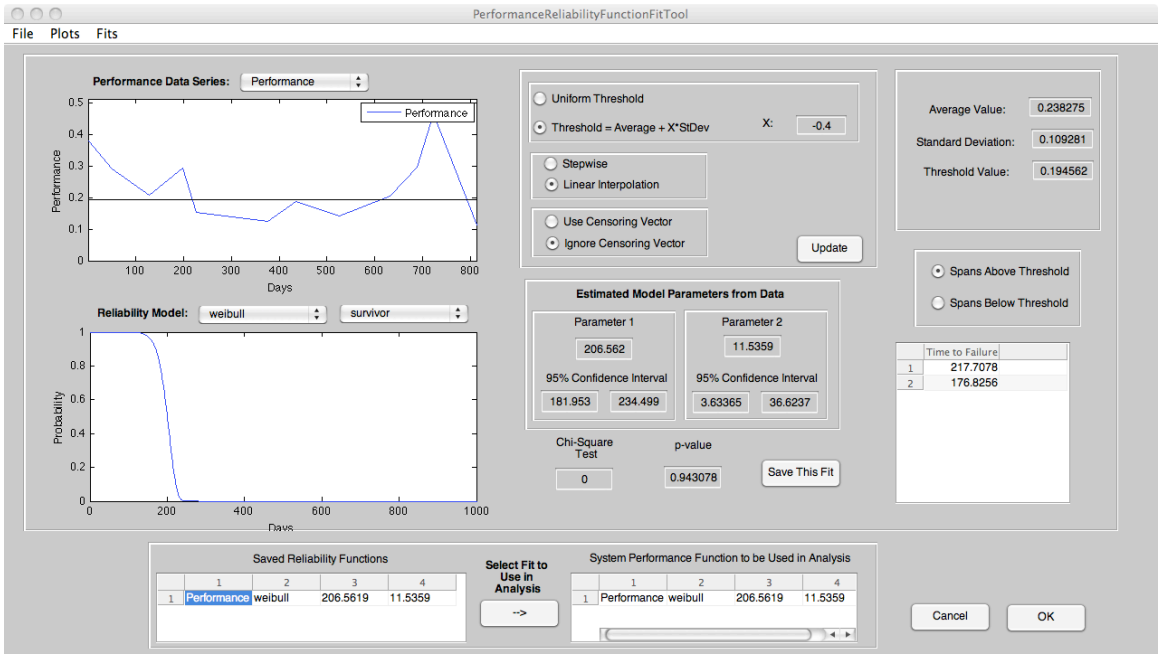


Figure 50. EcoReliAnT Performance Reliability Function Fit Tool is used to create a reliability function for the system based on user-defined performance data.

## 5. Analyze System Reliability

After defining a system reliability function and a model system, the fit of the model to the system performance reliability can be analyzed by clicking 'Analyze System'. Reliability functions of both the model and actual system will be displayed on the chart along with the sum of squared errors between the two curves in the box below the chart (Figure 51). This is calculated as the sum of squared errors between the two curves at intervals of 1 day.

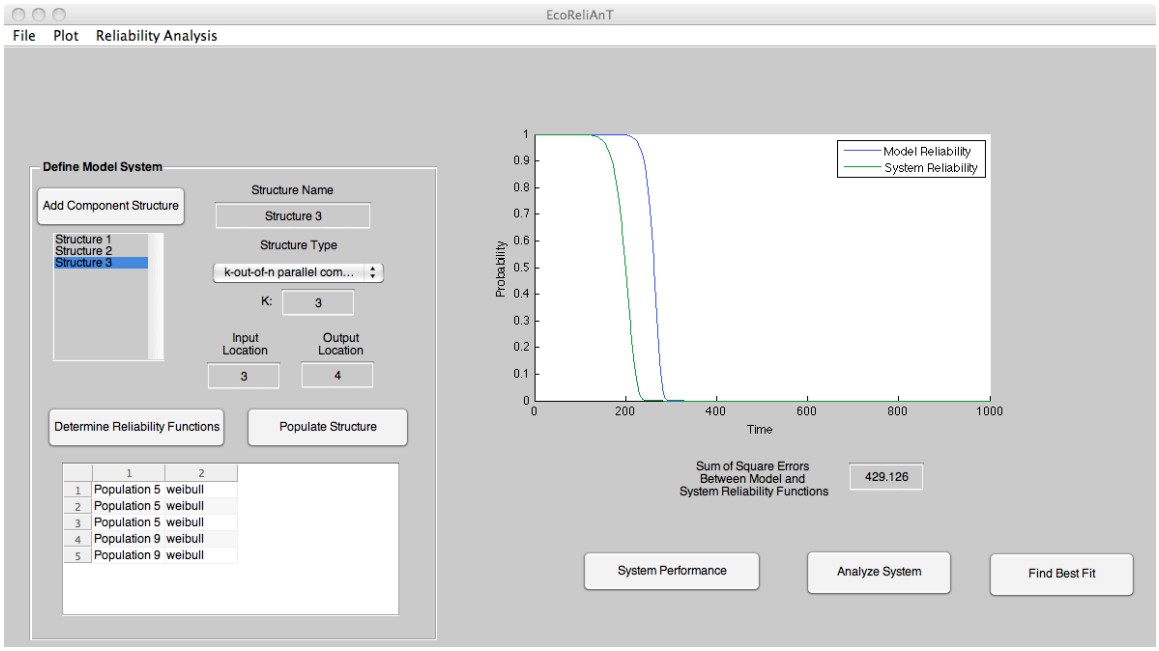



Figure 51. EcoReliAnT main window after defining and analyzing a model reliability system. The chart displays the reliability functions for the model and the real system based on performance data.

If there are any *k*-out-of-*n* systems in the model, clicking the 'Find Best Fit' button will determine the *k* values for all of the systems that will minimize the sum of squared errors between the two functions. With large system configurations containing many *k*-out-of-*n* systems, this optimization process can be time-consuming. Therefore, a status bar will pop up displaying the progress that has been made in the computation of the optimal reliability system.

After analyzing the reliability of the user-defined system, EcoReliAnT can determine which structure in the system is limiting overall system performance using the reliability importance of components. To display the reliability importance chart for each structure and select 'Reliability Importance of Components → Analyze by Structure' from the 'Reliability Analysis' menu. This will open the Reliability Importance of Components

window. In this window the reliability importance curve for a structure selected form the dropdown menu can be viewed (Figure 52).

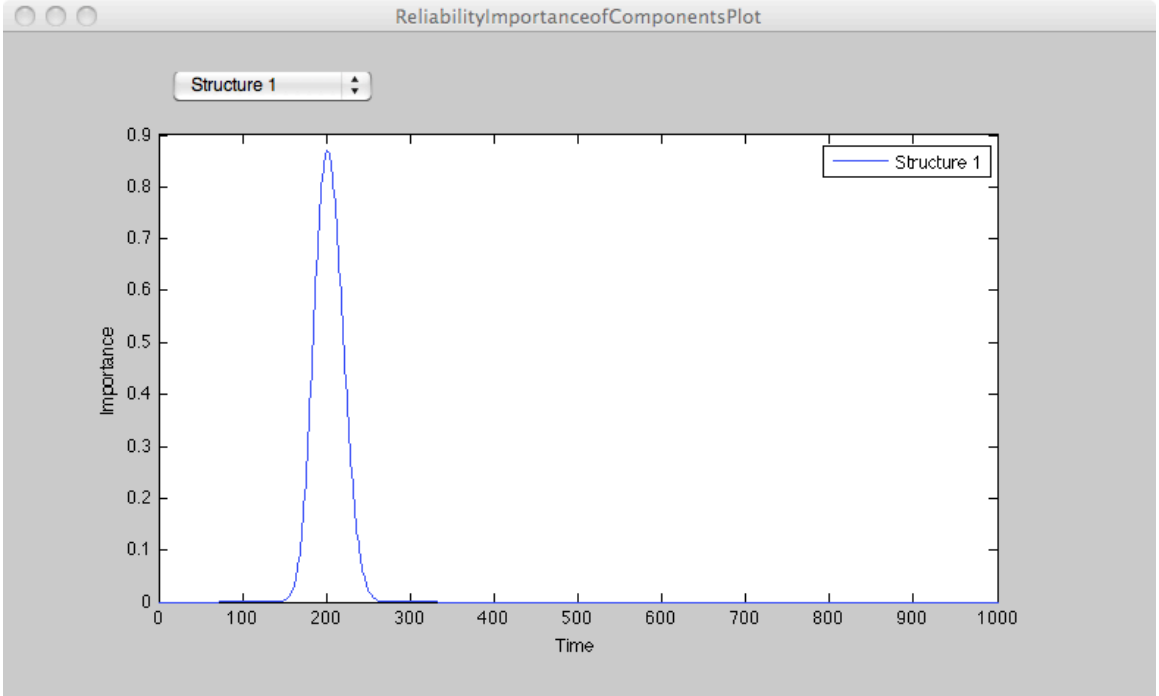



Figure 52. EcoReliAnT Reliability Importance of Components window displays the importance over time of the component structure selected from the dropdown menu.

The 'Limiting Structure' option in the dropdown menu will display a chart showing which of the component structures in the system analyzed has the largest importance value over time (Figure 53). As discussed earlier, an increase in the reliability of the component structure with the highest importance value will yield the largest increase in the system reliability, making it the reliability-limiting structure.

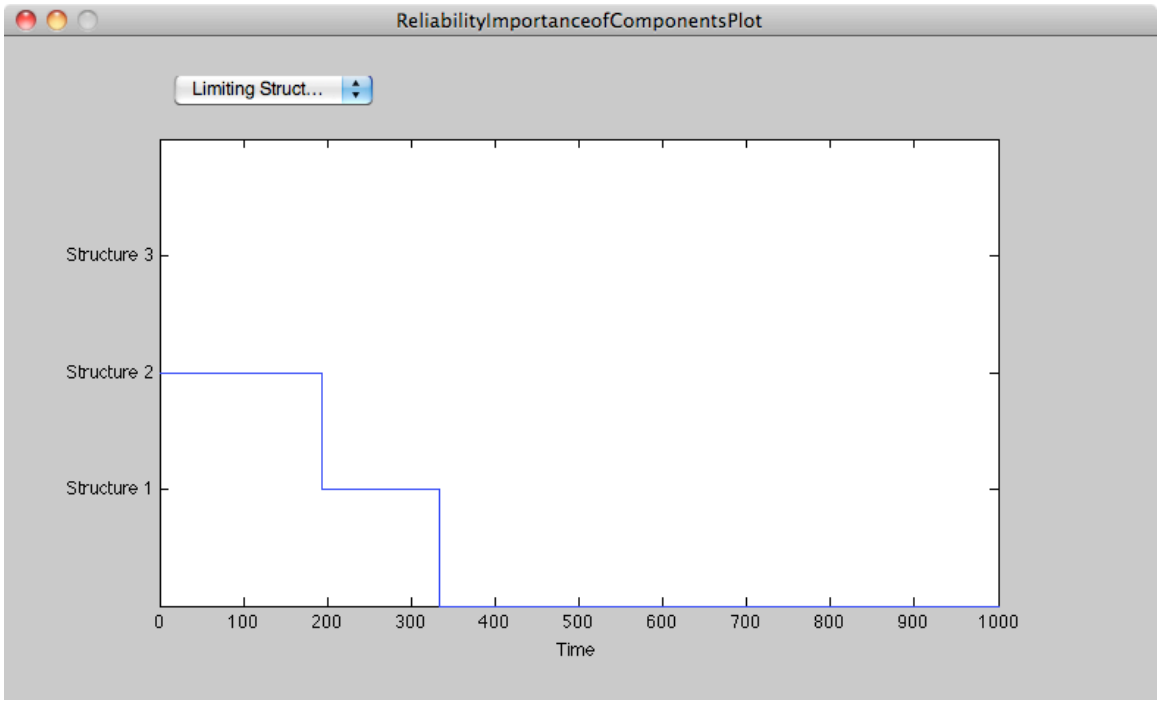



Figure 53. EcoReliAnT Reliability Importance of Components window displays a chart showing which component structure has the largest importance value over time when 'Limiting Structure' is selected from the dropdown menu.